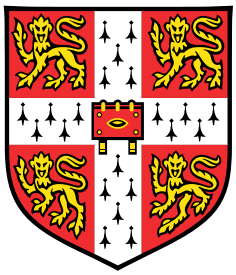
Drosophila melanogaster as a high-throughput and *in vivo* model to study body weight and feeding behaviour



Jennifer Chalmers

Pembroke College
University of Cambridge

This dissertation is submitted for the degree of
Doctor of Philosophy

October 2018

Declaration

I hereby declare that this dissertation is the result of my own work and includes nothing which is the outcome of work done in collaboration with others except as specified both here and in the text (work in section 4.5, and parts of figure 4.14, were done by Dr Loraine Tung).

It is not substantially the same as any that I have submitted, or, is being concurrently submitted for a degree or diploma or other qualification at the University of Cambridge or any other University or similar institution. I further state that no substantial part of my dissertation has already been submitted, or, is being concurrently submitted for any such degree, diploma or other qualification at the University of Cambridge or any other University or similar institution.

This dissertation contains fewer than 60,000 words excluding figures, photographs, tables, appendices and bibliography.

Jennifer Chalmers
October 2018

Acknowledgements

Firstly, I would like to express my sincere gratitude to my supervisors Dr Giles Yeo and Dr Loraine Tung for their never-ending support, patience, motivation, optimism, and knowledge. I could not have imagined having better or more inspirational advisors and mentors.

I thank my lab-mates in Bay 2 for the stimulating discussions, and for all the fun we have had in the last 3 years. Especially to Deb Rimmington and Kara Rainbow, for knowing the answer to any question I asked. My sincere thanks also goes to the Fly team - Dr Simon Collier, Glynnis Johnson, Boshka Gilanova, Kadri Oras, Lin Xi - who gave me access to their research facilities, and to Dr Che-Hsiung Liu for all of his advice. I also want to thank Dr Jenny Regan and Dr David Tree, for being the first to encourage me to pursue a PhD.

Thanks also to the Wellcome Trust and Pembroke College for providing the funding which enabled this research.

Last but not least, I would like to thank my family and friends for supporting me through this thesis.

Abstract

Obesity and its associated comorbidities represent one of the greatest public health challenges of the 21st century. Whilst the increase in prevalence has been driven by changes in lifestyle and environment, it is clear that there is a genetic component underlying the variation in body weight between individuals. However, the speed of translating obesity genetics into insightful biological knowledge has not kept pace with the rate of identification of potential genes of interest. One reason for this is that studies have so far been addressed in complex model organisms such as mice.

Drosophila melanogaster is a key model for research in developmental biology, cell biology and neurobiology, and has also recently been demonstrated to be an excellent model for dissecting metabolic homeostasis pathways. In this thesis, a suite of assays was developed in *Drosophila* to examine changes in the state of energy homeostasis and changes in energy intake in response to specific genetic perturbations. This screen was validated by its ability to detect metabolic perturbation in wild-type flies, and to differentiate between positive and negative control genes. Different *Drosophila* models were investigated for their suitability for the screen, and neuron-specific RNAi was found to provide the most useful information. CRISPR was also investigated as a method of generating mutations in target genes *de novo*.

The validated screen was then used to explore genes highlighted by GWAS studies of human body mass index (BMI), and by transcriptomics studies of fasted mouse hypothalamic neurons. Each of these screens identified some genes which are already known to play an important role in the control of energy homeostasis in mammals, demonstrating that the screen is able to produce relevant results. They also highlighted several genes for which there is currently no published functional data, demonstrating the utility of *Drosophila* in a high-throughput screen to assay the potential involvement of genes in the neuronal control of food intake and body weight.

Table of contents

Abbreviations

Roman Symbols

AKH	Adipokinetic Hormone
APC	AKH Producing Cell
ARC	Arcuate nucleus
BMI	Body Mass Index
bp	Base Pairs
CAFE	Capillary Feeder
CNS	Central Nervous System
CRISPR	Clustered Regularly Interspaced Short Palindromic Repeats
CyO	Chromosome 2 balancer
dILP	*Drosophila* Insulin Like Peptide
DMN	Dorsomedial Nucleus
DSB	Double Strand Break
FACS	Fluorescence Activated Cell Sorting
FM7	Chromosome 1 balancer
GFP	Green Fluorescent Protein
GI	Gastrointestinal
gRNA	Guide RNA

GWAS	Genome Wide Association Study
HFD	High Fat Diet
HSD	High Sucrose Diet
HSL	Hormone Sensitive Lipase
ICV	Intracerebroventricular
IMPC	International Mouse Phenotyping Consortium
IPC	Insulin Producing Cell
KO	Knock out
LD	Lipid Droplet
LH	Lateral Hypothalamus
Lk	Leucokinin
Lkr	Leucokinin receptor
LoF	Loss-of-Function
ND	Normal Diet
NHEJ	Non Homologous End Joining
NmU	Neuromedin U
NPF	Neuropeptide F
NPK	Neuropeptide K
NPY	Neuropeptide Y
PAM	Protospacer Associated Motif
PBS	Phosphate buffered saline
POMC	Proopiomelanocortin
PVN	Paraventricular Nucleus
RNAi	RNA interference

siRNA	Small Interfering RNA
SM	Starvation Media
SNP	Single Nucleotide Polymorphism
sNPF	Short Neuropeptide F
TAG	Triacylglycerol
TM6b	Chromosome 3 balancer
UAS_G	Upstream Activating Sequence
UTR	Untranslated region
VDRC	Vienna Drosophila Resource Center
VMN	Ventromedial Nucleus
YFP	Yellow Fluorescent Protein

Chapter 1

Introduction

1.1 Obesity and public health

Being overweight is defined as having an excessive accumulation of fat that negatively impacts health. In practice, body fat is often assessed using Body Mass Index (BMI - a person's weight divided by the square of their height): $>$25kg/m^2 is classed as overweight and $>$30kg/m^2 as obese. Average body weight has been increasing rapidly all over the world and obesity is now classed as a global epidemic by the World Health Organization. Since 1975, the proportion of obese people has more than tripled and today 39% of adults are overweight and an additional 13% are obese, which is significantly more than the $<$10% who are underweight (GBD, 2017; NCD Risk Factor Collaboration, 2016; WHO, 2018).

The high levels of body fat in obesity increases the risk of many diseases including type 2 diabetes, cardiovascular disease, hypertension, cancer, and mental health (Must et al., 1999). Therefore, high body weight is a major contributor to the global burden of disease. High BMI is attributed to 4 million deaths each year globally, and 14% of premature deaths in Europe (GBD, 2017; Di Angelantonio et al., 2018). Obesity also has an economic cost. For example in 2007 the cost to the NHS of treating overweight and obese patients was £5.1 billion and an additional £11.6 billion was spent on costs related to consequent unemployment (Scarborough et al., 2011).

Over recent years, governments and health professionals have made many attempts to reduce obesity but these have proved unsuccessful and the prevalence is still rising. It is predicted that $>$20% of the world's population will be obese by 2025 (NCD Risk Factor Collaboration, 2016). Gaining a full understanding of the causes of obesity is therefore critical in order to find effective and sustainable solutions.

1.2 Environmental factors influencing obesity

An individual's body weight is a reflection of the balance between their energy intake (food and drink ingested) versus their energy output (for example movement, heat generation, immune system function). A stable weight is maintained when these two figures are equal but any discrepancies cause weight gain or weight loss. Large prolonged excesses of energy intake lead to obesity and metabolic disease.

Because it has only become a problem relatively recently, the obesity epidemic is mainly attributed to our modern lifestyle. Calorie intake has risen due to the increased abundance of easy-to-access and processed high-calorie food whereas energy expenditure has decreased as both work and leisure have become more sedentary (French et al., 2001). Together these factors create an obesogenic environment which tips the energy balance towards weight gain.

Although lifestyle changes are undoubtedly responsible for the recent surge in obesity, there is also evidence that other factors influence obesity risk such as *in utero* exposure to maternal nutrition, sleep patterns, the microbiome, and obesity in social acquaintances (Rooney and Ozanne, 2011; Hruby and Hu, 2015). In addition, genetics has been shown to be an important factor for determining obesity risk.

1.3 Genetic factors influencing obesity

The genes and alleles present in the population have not changed significantly in the same time frame as the obesity epidemic has developed. Instead, the environment has become obesogenic and how an individual responds to this environment is influenced by their genetics. The earliest evidence for a role of genetics in obesity risk came from studying families. For example, fat mass is more similar between monozygotic twins than it is between dizygotic twins (Stunkard et al., 1986a; Feinleib et al., 1977), even when they are raised apart (Stunkard et al., 1990). Further, the BMI of adopted individuals correlates with their biological parents and not with their adoptive parents (Stunkard et al., 1986b). These and similar studies have together concluded that up to 70% of the variation in BMI and fat mass between individuals is due to genetic factors (Maes et al., 1997).

Monogenic obesity

A monogenic disorder is a disease caused by a deleterious mutation in a single gene. For such disorders, genetics has a strong impact with only a small influence from the environment. Monogenic obesity is rare, severe and early-onset (Huvenne et al., 2016) and mutations in several different genes are known to cause this. One example is leptin: congenital leptin deficiency is inherited in an autosomal recessive manner and results in intense hyperphagia, leading to severe obesity from early childhood (Montague et al., 1997; Mutch and Clement, 2006; Echwald et al., 1997; Oksanen et al., 1997). Studies of leptin-deficient patients and rodent models have shown that leptin is a hormone released by adipose tissue as a signal of the level of triacylglycerol (TAG) stores (Considine et al., 1996; Halaas et al., 1995). The hyperphagia (and other symptoms such as infertility) seen in leptin deficient individuals are due to the inability of their adipose tissue to communicate to the body that there is fat stored.

Fewer than 100 patients worldwide are known to have homozygous deleterious mutations in their leptin gene (Huvenne et al., 2016) and in total, monogenic obesity syndromes account for $<5\%$ of obesity cases (Xia and Grant, 2013). Nevertheless, much of our knowledge about the physiological systems involved in energy homeostasis has come from identifying humans with monogenic obesity syndromes and studying genetically engineered rodent models. Most of the known monogenic obesity mutations are in genes that are part of the leptin-melanocortin axis which has a key role in the regulation of food intake (see section 1.4) for example leptin (Zhang et al., 1994; Montague et al., 1997), the leptin receptor LEPR (Tartaglia et al., 1995; Chua et al., 1996), POMC (Challis et al., 2004; Krude et al., 1998), and PCSK1 (Jackson et al., 1997).

Oligogenic obesity

2-3% of obese adults and children have a genetic mutation which produces an obesity phenotype of variable severity which is partly dependent on the environment, but which has no additional specific phenotypes (Huvenne et al., 2016). This is referred to as oligogenic obesity. An example is the MC4R gene for which mutations have an autosomal dominant mode of transmission, incomplete age-related penetrance, and variability in the severity of hyperphagia and obesity (MacKenzie, 2006).

Syndromic obesity

There are also more than 100 syndromic disorders which are associated with severe obesity. Unlike monogenic and oligogenic obesity disorders, the syndromic obesity syndromes have additional distinct clinical phenotypes, often mental retardation, dysmorphic features, and/or developmental abnormalities (Huvenne et al., 2016). Whilst other syndromes that affect cognition such as Down syndrome show an increased incidence of obesity, syndromic obesity has specific effects on food intake (Chung, 2012).

The most common syndromic obesity disorder is Prader-Willi Syndrome which affects 1 in 15,000-25,000 live births (Chung, 2012) and is caused by a deletion of the region on paternal chromosome 15 which encompasses the SNORD116 locus (Bieth et al., 2015). In recent years genetic sequencing of patients with syndromic obesity disorders has identified mutations in genes such as SH2B1 (Doche et al., 2012), TUB (Borman et al., 2014), CPE (Alsters et al., 2015), NTRK2 (Yeo et al., 2004), SIM1 (Michaud et al., 2001), ACP1, TMEM18, MYT1L (Doco-Fenzy et al., 2014), KSR2 (Pearce et al., 2013), and RAI1 (Alsters et al., 2015).

Polygenic obesity

Monogenic, oligogenic and syndromic obesity disorders are rare. Instead the obesity in most patients has a polygenic aetiology and is referred to as common obesity. This type of obesity is influenced by the combined effect of common subtle variants in multiple different genes. Individually these genetic variants only have a small effect, but they interact with each other and with the environment. Given that the control of body weight has likely been subject to a number of evolutionary selection pressures, it is not surprising that common obesity is a complex polygenic trait. It is only with the advent of modern genetic techniques that the genes involved in polygenic obesity are starting to be revealed.

Early studies of common obesity looked for variants in genes that had already been identified in monogenic obesity. For example, common variants in the leptin gene are associated with BMI (Jiang et al., 2004; Li et al., 1999; Mizuta et al., 2008). However, because monogenic obesity is rare, the number of genes that can be identified in this way is limited. Further, any genes for which severe mutations cause lethality will never be found. Therefore, hypothesis-driven approaches have only explained a small fraction of the genetic risk for obesity.

The leading non-hypothesis-driven approach is currently a Genome Wide Association Study (GWAS). GWASs genotype millions of single nucleotide polymorphisms (SNPs) in thousands of people. The SNPs are spread across the entire genome in an unbiased and dense manner. The frequency of a variant is compared between disease and control subjects and any significant differences signal the presence of as allele that confers risk or protection.

Since 2007, multiple GWASs have been performed for several different obesity-related characteristics such as BMI, body fat percentage and waist-to-hip ratio. These studies have identified new genetic loci associated with obesity that have then been replicated across many large independent studies with hundreds of thousands of subjects. The effect sizes of these GWAS SNPs are much smaller than those seen in monogenic obesity - for example the largest effect is seen with the FTO locus where homozygotes for the rs9939609 risk allele weigh 3kg more and have a 1.67-fold increased risk of obesity than homozygotes for the protective allele (Frayling et al., 2007). However, these genetic variants are much more common than the mutations associated with monogenic obesity which arguably makes their impact greater - for example 18% of people with European ancestry carry 2 risk FTO alleles and 46% carry 1 (Frayling et al., 2007).

GWASs have associated BMI with SNPs near many of the genes involved in monogenic, oligogenic and syndromic obesity, for example LEPR, POMC, MC4R, PCSK1, SH2B1, TMEM18, BDNF, and TUB, as well as >100 novel genetic loci (Speliotes et al., 2010; Locke et al., 2015; Akiyama et al., 2017).

1.4 Neuronal control of feeding behaviour

Although the obesity epidemic is caused / influenced by a large number of factors, weight gain can *only* occur when energy input from feeding exceeds energy output. Therefore the study of feeding behaviour (energy intake) is an important part of understanding the regulation of energy homeostasis and its dysregulation in obesity.

Feeding is a fundamental activity of all animals as they require energy for growth, survival and reproduction. Therefore species have evolved feeding habits based on their internal metabolic needs and external sensory signals which together affect meal size, frequency, and content.

The central nervous system (CNS) is key for processing relevant information to generate appropriate feeding responses. In particular, the hypothalamus is a critical mediator of energy homeostasis in mammals. Early studies showed that introducing lesions into specific regions of the rat hypothalamus caused hyperphagia and weight gain (Hetherington and Ranson, 1942). Studies of monogenic, oligogenic and syndromic obesity have provided insight into the underlying molecular mechanisms that control food intake in the CNS.

The best characterised pathway is the leptin/melancortin axis in the hypothalamus which is summarised in figure 1.1. After release from the adipose tissue, leptin travels in the blood to its target tissues. The arcuate nucleus (ARC) of the hypothalamus is located near the semipermeable blood brain barrier, allowing it to sense and respond to peripheral signals of energy state such as leptin. In the ARC, key populations of neurons express the leptin receptor (Fei et al., 1997). One of these sets of neurons also express POMC and are thus referred to as POMC neurons. Binding of leptin to its receptor on these cells causes depolarisation of the neurons (Cowley et al., 2001). It also activates the JAK-STAT signalling pathway which includes SH2B1 and TUB (Ren et al., 2005; Carroll et al., 2004) and results in increased expression of POMC (Ernst et al., 2009; Yazdi et al., 2015). The POMC precursor peptide is cleaved into several proteins including α-MSH and β-MSH by the enzymes PCSK1 and CPE (Bertagna, 1994; Castro and Morrison, 1997). α-MSH and β-MSH are secreted by active POMC neurons and can then bind to MC4R on the surface of downstream neurons. Various genes are known to be involved in signalling pathways in the downstream neurons including BDNF and its receptor TrkB (Xu et al., 2003; Kernie et al., 2000; Yeo et al., 2004; Gray et al., 2007).

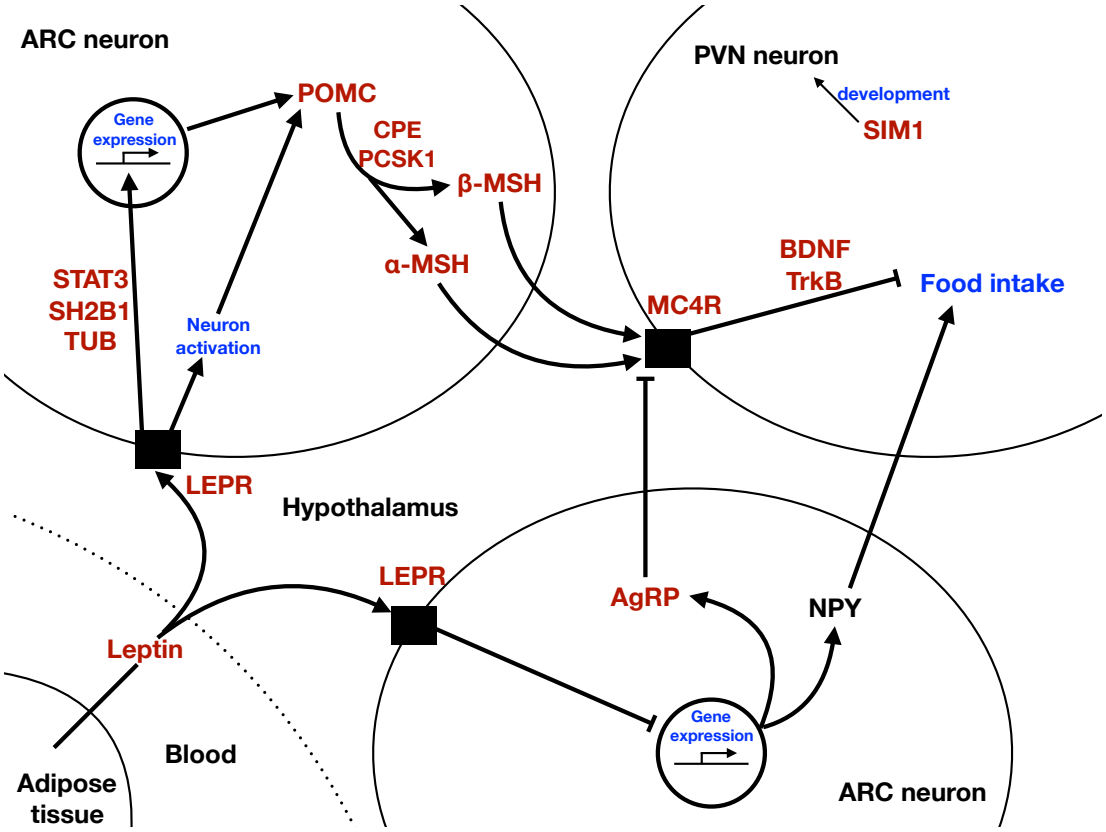


Fig. 1.1 The leptin-melanocortin signalling pathway influences food intake. Highlighted in red are genes found in monogenic/oligogenic/syndromic obesity.
AgRP = agouti-related peptide, ARC = arcuate nucleus, BDNF = brain-derived neurotrophic factor, CPE = carboxypeptidase E, LEPR = leptin receptor, MC4R = melanocortin 4 receptor, α/β-MSH = α/β-melanocyte-stimulating hormone, NPY = neuropeptide Y, PCSK1 = proprotein Convertase Subtilisin/Kexin Type 1, POMC = pro-opiomelanocortin, PVN = paraventricular nucleus, SH2B1 = SH2B adapter protein 1, SIM1 = single-minded homolog 1, STAT3 = signal transducer and activator of transcription 3, TrkB = tropomyosin receptor kinase B, TUB = tubby Bipartite Transcription Factor.

The POMC neurons in the ARC project to other areas of the hypothalamus including the ventromedial nucleus (VMN), dorsomedial nucleus (DMN), paraventricular nucleus (PVN) and lateral hypothalamus (LH) (see figure 1.2) (Cone, 2005). Studies in which these hypothalamic areas are individually ablated, genetically perturbed, or injected with specific neuropeptides have shown that they are all involved in feeding behaviour regulation (Simpson et al., 2009). POMC neurons also project to other areas of the brain, for example the amygdala which is involved in reward. By signalling to all of these areas, the leptin-melanocortin pathway inhibits food intake (anorexigenic) and so it is unsurprising that mutations in many of the involved genes have been found in obesity syndromes.

Not all of the genes identified from obese patients are directly part of the leptin-melanocortin pathway, but many still act in a related manner. For example SIM1 encodes a transcription factor that is involved in development of the PVN (Michaud et al., 2001).

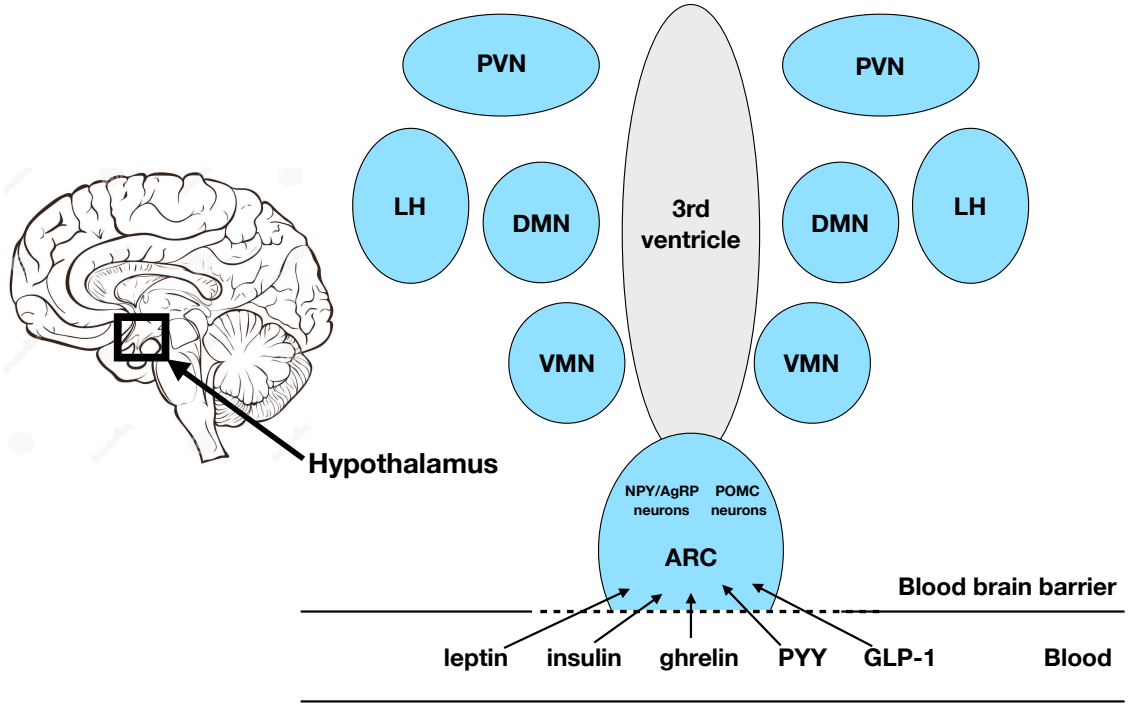


Fig. 1.2 The main hypothalamic nuclei involved in the regulation of appetite in the human brain. Diagram based on Simpson et al. (2009).
ARC = arcuate nucleus, DMN = dorsomedial nucleus, GLP1 = glucagon-like peptide-1, LH = lateral hypothalamus, PVN = paraventricular nucleus, PYY = peptide YY, VMN = ventromedial nucleus.

Another population of neurons in the ARC also sense and respond to leptin: the NPY/AgRP neurons (figure 1.1). Leptin binding to LEPR on these neurons inhibits the production of both NPY and AgRP proteins (Varela and Horvath, 2012). Release of NPY and GABA from these neurons to receptors in the PVN, DMN, VMN and LH has an orexigenic (appetite-stimulating) effect, and release of AgRP antagonises MC4R to prevent its anorexigenic effect (Ollmann et al., 1997). Consequently, constitutive expression of AgRP in mice results in hyperphagia and obesity (Graham et al., 1997).

As well as leptin, the POMC and NPY/AgRP neurons in the ARC respond in a similar manner to other circulating signals of peripheral energy surplus including insulin and the post-prandial gut peptides PYY and GLP1 (Brüning et al., 2000; Teubner and Bartness, 2013; NamKoong et al., 2017). The ARC neurons respond in the opposite manner to ghrelin, a hormone released by the stomach during fasting (Riediger et al., 2003).

Finally, GWAS studies of common obesity have also shown the importance of the brain for regulating energy homeostasis: the genes in loci which correlate with BMI (a measure of total fat mass) are enriched for expression in the CNS (Locke et al., 2015). By comparison, genes associated with waist-to-hip ratio (a measure of the distribution of body fat) are instead enriched in adipose tissue (Shungin et al., 2015).

Therefore, studying the CNS is key to understanding mechanisms of energy homeostasis and its dysregulation in obesity.

1.5 Why study obesity genetics?

Discovery of genes which influence common obesity make it possible to identify those who are at risk, so that lifestyle modifications (changed diet and increased exercise) can be made before weight gain and its comorbidities even occur. However, for those already suffering from obesity there is currently no treatment that is both safe and highly effective. Lifestyle modifications are the safest means of weight loss, but patient compliance is generally poor which reduces efficacy. Bariatric surgery can result in weight loss of 20%, but like all surgery it is expensive and risky. Currently five pharmaceutical agents are approved for use in the US for weight loss and one in the UK, but these drugs have major side-effects and only bring about modest weight loss that is not maintained longer than one year (Gadde et al., 2018).

The discovery of leptin led to the development of an effective treatment for leptin-deficient patients: subcutaneous injection of leptin reduces hyperphagia and food intake and leads to weight loss (Farooqi et al., 1999). Another drug currently in clinical trials for patients with deficiencies of LEPR or POMC is setmelanotide: an agonist of MC4R shown to decrease food intake and cause weight loss in humans (Clément et al., 2018; Ayers et al., 2018). These two treatments show how a genetic discovery can directly lead to development of a novel pharmaceutical reagent for the treatment of obesity. However, leptin injection is not effective for the treatment of common obesity (Heymsfield et al., 1999) because most of these individuals have elevated levels of leptin and are instead leptin resistant (Frederich et al., 1995).

Therefore understanding the biology of pathways and systems that are involved in energy homeostasis is necessary in order to develop treatments for common obesity. Studying the function of individual genes using single gene mutants is a good way to dissect the genetic basis of a disease pathway (Moore, 2010). Since genetics strongly impacts body weight, genetic approaches offer an effective tool for identifying and understanding the complex molecular mechanisms of energy homeostasis. The polygenic basis of common obesity makes it a rich source of potential novel insights which may, in the future, aid development of pharmaceutical agents which are both effective and specific.

1.6 The difficulties of studying obesity genetics

Following completion of the human genome project, the advent of large-scale sequencing methods has driven the study of human disease genetics forward at a remarkable pace. Genomic techniques have identified hundreds of human genes associated with BMI, fat mass, and other obesity-related characteristics. Examples of such techniques include GWASs and RNAseq.

The problem is that because the control of energy homeostasis involves large numbers of genes, interactions between these genes and interactions with the environment, these genetic techniques generate large data sets. Unfortunately the scientific community lacks the resources to study all of the genes identified in detail and so few have been followed up functionally.

As techniques become more advanced, results are being generated at an ever increasing rate. The question then arises as to which of the identified genes are likely to be the most fruitful to study. What is needed is a method of prioritising the genes from these large data sets to determine which should be the subject of further research.

1.7 Animal models for the study of obesity genetics

To study obesity genetics relevant to humans, the most accurate model would, of course, be humans. However the obvious ethical barriers to creating mutants *de novo* mean that research relies instead on the serendipitous identification of rare genetic mutants. Even then, the person's environment is not controlled which decreases the power to link genetic variation to phenotype. Further, it is difficult to obtain tissue samples from the brain (the most relevant tissue).

To date, most obesity genetics studies have used mice and rats. These rodents share 99% of their genome with humans as well as many biochemical pathways, physiological features and pathological symptoms (Vandamme, 2014; Rosenthal and Brown, 2007). However, rodents have several characteristics which make them unsuitable for looking at large genetic data sets. Firstly, experiments are time-consuming with low statistical power because breeding rodents is slow (generation time of 10 weeks with $<$10 pups per litter (Phifer-Rixey and Nachman, 2015)). Experiments are also expensive due to the high cost of purchasing (>£15 for 1 wild-type from the Jackson Laboratory) and maintaining (>£5.60 per week per mouse in the Cambridge University PDN mouse facility) the animals. The estimated cost of producing and phenotyping a mouse model for a single gene is £30,000 (Moore, 2010).

Cell culture has also provided many useful insights into obesity genetics, and unlike with rodents, such experiments are fast and cheap. However, their unicellularity means that cells can have only limited use for studying a systems disorder such as obesity or related behaviours like feeding.

Drosophila melanogaster have been used as a model organism for over 100 years. They have many features that make them potentially well suited for use in high-throughput genetic studies of energy homeostasis. These attributes will now be explored.

1.8 *Drosophila* as a high-throughput model

It is economical to work with *Drosophila*: stocks are freely shared amongst researchers or purchased cheaply (£1-£2) from international stock centres and it is 10,000 times cheaper to maintain fly stocks than mice (Neckameyer and Argue, 2013). Experiments using *Drosophila* are also time-efficient: the generation time is only 10 days and 1 female fly can lay up to 100 embryos each day (Jennings, 2011). These two factors make experiments with *Drosophila* high-throughput and thus suitable for studying large numbers of genes.

For studying genetics it is important to be able to generate specific genotypes. In this regard, flies offer the following advantages: (i) there are only four chromosome pairs to consider, (ii) there is no meiotic recombination in males, (iii) it is easy to distinguish males vs females and virgins vs mated flies, and (iv) transgenes can be tracked with many different markers that affect external characteristics visible using just a bench microscope.

Further, a huge variety of genetic techniques have been developed for use in *Drosophila*. A plethora of mutant and transgenic *Drosophila* lines have been created which are readily available and enable over-expression, mis-expression, knock-down, or knock-out (KO) of virtually any gene in a specific tissue or condition. Four genetic techniques are used in this thesis, and will be explained in more detail in chapters 3 and 5: transposon insertion, the GAL4-UAS system, RNA-interference (RNAi), and CRISPR.

Another important factor for genetic studies is that only 25% of *Drosophila* genes are adult-lethal when mutant, and many of these are still viable as larvae. Even mutants for most housekeeping genes are able to complete embryogenesis (Wangler et al., 2017). This means that is is possible to study most genes in a homozygous mutant fashion. For those genes which are not viable, many tissues in *Drosophila* are amenable to mosaic analysis which allows lethal genes to be perturbed and studied in just a single tissue which is often less detrimental than mutation throughout the entire fly. Alternatively, balancer chromosomes are readily available and allow the mutations to be maintained in heterozygote populations. Balancer chromosomes are engineered constructs with DNA inversions to suppress meiotic recombination (Hentges and Justice, 2004), and since they are homozygous lethal/sterile and carry dominant phenotypic markers that can be scored visually, there is no need to constantly PCR genotype as is the case with mice.

1.9 Genetic conservation between *Drosophila* and humans

Drosophila have been used extensively in the past as a model for human disease and this is possible because of high levels of genetic conservation. Studies have shown that 53% of human genes are present in *Drosophila*, but genes linked to human disease show even higher conservation with 77% having an orthologue in flies (Fortini et al., 2000; Reiter et al., 2001; Wangler et al., 2017). The nucleotide/protein sequence identity of homologues is usually around 40% but can be more than 90% within functional domains (Pandey and Nichols, 2011).

As well as sequence conservation there is also conservation of tissue expression patterns. For example, of human genes associated with neurological disorders, more fly homologues were found to be expressed in the fly CNS than in the fly digestive system (Wangler et al., 2017).

Even for genes with no apparent sequence homologue, there are studies in which expressing the human gene in flies has yielded useful information (St Johnston, 2002). For example, expression of human synphilin-1 in fly neurons increases food intake, body weight, and fat deposition which is the same phenotype observed in mice (Li et al., 2012; Liu et al., 2012; Smith et al., 2014). Further, mutations in *upd1* increase food intake and TAG stores in flies but these phenotypes can be rescued by expression of the human leptin gene (Beshel et al., 2017).

Importantly for this thesis, energy homeostasis in *Drosophila* is also influenced by genetics. Several genes have been found which model monogenic obesity in *Drosophila* and some examples such as *upd1* will be discussed in section 1.11. Further, applying starvation as an evolutionary selective pressure over 15 generations (only those that survived starvation were mated) eventually resulted in flies with hugely increased stores of TAG and sugar, and reduced activity levels compared to controls (Schwasinger-Schmidt et al., 2012). It is also possible to perform GWASs on flies to identify SNPs that are associated with variations in metabolic traits, for example Jehrke et al. (2018).

1.10 Metabolic conservation in *Drosophila*

Many physiological and biochemical pathways involved in energy storage and utilisation are well conserved across the animal kingdom, including between humans and *Drosophila*. This is an important point if flies are to be used to study energy homeostasis, and four examples will be shown briefly here.

1.10.1 Conserved metabolic tissues and physiology

As defined by their structure, function, and gene expression patterns, all of the principal mammalian metabolic tissues and organs have homologues in *Drosophila*, as shown in the table below and figure 1.3. Like in mammals, the regulation of energy homeostasis in *Drosophila* requires the interplay between all of these metabolically active tissues in changing environmental conditions.

Mammalian organ	*Drosophila* organ	Function
adipose tissue	fat body	stores TAG, secretes adipokine-like factors
liver	fat body	stores and regulates glycogen
liver	oenocytes	stores and regulates lipids, senses metabolic status
GI tract	GI tract	digestion and absorption of nutrients
kidneys	malpighian tubules, nephrocytes	excretion, osmoregulation, detoxification
pancreas	Insulin-like peptide producing cells (IPCs)	glucose homeostasis
	Adipokinetic Hormone (AKH) producing cells (APCs)	
blood	haemolymph	circulates nutrients and hormones

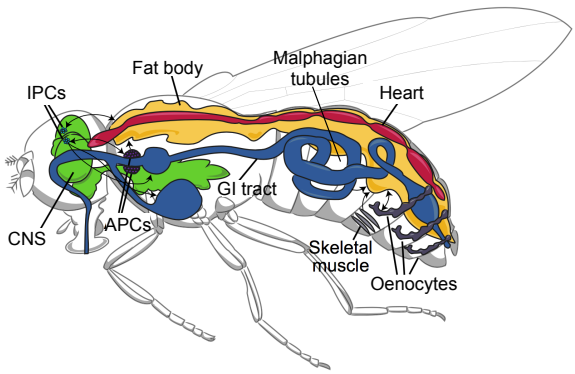


Fig. 1.3 Metabolic organs in *Drosophila*. Figure based on Musselman and Kühnlein (2018).

1.10.2 Conserved metabolic pathway: lipid storage

In times of energy excess, both mammals and *Drosophila* store TAG in the form of lipid droplets (LDs) in adipose tissue and the fat body, respectively. This TAG is released into circulation during periods of scarcity.

Formation of TAG

The stored TAG is formed by the Kennedy pathway in four enzymatic steps (Yen et al., 2008). The first and second steps are largely unexplored in *Drosophila* (Kühnlein, 2012). The third step is catalysed by the highly conserved lipin gene family: Lipin-1 deficient mice exhibit lipodystrophy and insulin resistance (Peterfy et al., 2001; Reue, 2009) and *Drosophila* larvae lacking *lipin* have smaller LDs and a smaller fat body (Ugrankar et al., 2011). The final step is catalysed by diglyceride-acyltransferase enzymes: DGAT1-deficient mice have 50% less adiposity and are resistant to diet-induced obesity (Smith et al., 2000) and adult *Drosophila* with knockout of *mdy* are also lean (Beller et al., 2010).

Basal lipolysis

Under basal conditions, mammalian perilipin and *Drosophila* Lsd2 proteins are bound to the surface of LDs where they prevent access by lipases (Wolins et al., 2006). Consequently, humans, mice and flies which lack functional perilipin/Lsd2 have smaller TAG stores due to increased basal lypolysis (Gandotra et al., 2011; Tansey et al., 2001; Teixeira et al., 2003). Conversely, overexpression of *Lsd2* in the fat body of flies increases fat accumulation (Arrese et al., 2008; Fauny et al., 2005; Gronke et al., 2003). Basal lipolysis is mediated by ATGL in mammals and bmm in *Drosophila*, and mutants in these genes show excessive accumulation of lipids (Gronke et al., 2005; Haemmerle et al., 2006; Schweiger et al., 2009).

TAG release

In times of negative energy balance, glucagon / adipokinetic hormone (AKH) bind to their receptors on adipose / fat body cells, which activates PKA. In turn, PKA phosphorylates perilipin-1 / Dmplin1 proteins which makes them able to recruit hormone-sensitive lipase (HSL) to the LD, triggering lipolysis (Arrese et al., 1999, 2008; Bi et al., 2012; Miyoshi et al., 2006; Sztalryd et al., 2003; Kühnlein, 2012). Mouse cells with loss of perilipin-1 phosphorylation can not enhance lipolysis (Miyoshi et al., 2006), and starvation-induced

lipolysis is blunted in *AKHR* mutant flies whereas ectopic expression of *AKH* in the fat body decreases stored lipids (Gronke et al., 2007; Lee and Park, 2004; Kühnlein, 2012).

1.10.3 Conserved metabolic signalling: glucose homeostasis

Insulin release

In both mammals and flies, carbohydrate homeostasis is controlled by hormones secreted into the blood and haemolymph circulation respectively. In response to high glucose levels, mammalian pancreatic β cells secrete insulin and *Drosophila* insulin-producing cells (IPCs) secrete insulin-like peptide (dILP)-2, -3 and -5 (Brogiolo et al., 2001; Fridell et al., 2009; Kreneisz et al., 2010). Consequently, ablation of pancreatic β cells / IPCs or mutation of insulin / *dilp1-5* results in elevated levels of sugars in the blood / haemolymph (Atkinson and Maclaren, 1994; Nishi and Nanjo, 2011; Rulifson et al., 2002; Zhang et al., 2009). The development of IPCs is controlled by the transcription factor *eyeless* whose mammalian homologue, Pax6, is required for β-cell specification (Clements et al., 2008).

In β cells, glucose sensing by GLUT1 and/or GLUT2 triggers release of ATP from mitochondria (Graham and Pick, 2017). GLUT1 is also involved in flies as knockdown of *glut1* in IPCs decreases levels of circulating dILP2 (Park et al., 2014).

In mammals, the released ATP regulates K^{ATP} channels in the cell membrane, causing depolarisation which in turn activates Ca^{2+} channels, leading to exocytosis of insulin from β cells (Nassel and Broeck, 2016). In *Drosophila*, K^{ATP} channels are also involved, as shown using glibenclamide (a K^{ATP} channel inhibitor) and glucose causes an influx of Ca^{2+} in IPCs (Fridell et al., 2009; Kreneisz et al., 2010).

Insulin action

Insulin / dILPs travel in the blood / haemolymph to their target tissues and bind to the structurally conserved insulin receptor. This initiates a conserved signalling pathway shown in figure 1.4.

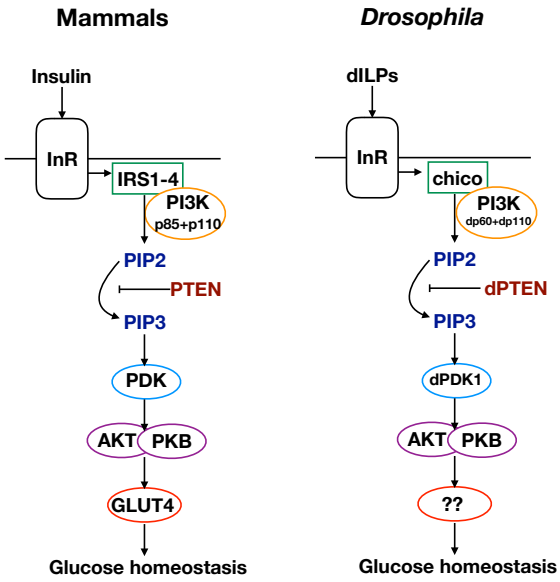


Fig. 1.4 The insulin signalling pathway is conserved between mammals and *Drosophila*. dILP = *Drosophila* insulin-like peptide, GLUT4 = glucose transporter 4, InR = insulin/ILP receptor, IRS = insulin receptor substrate 1, PDK = phosphoinositide-dependent kinase-1, PIP2 = phosphatidylinositol 4,5-bisphosphate, PIP3 = phosphatidylinositol 3,4,5-trisphosphate, PI3K = phosphoinositide 3-kinase, AKT/PKB = protein kinase B, PTEN = phosphatase and tensin homolog.

This signalling pathway activates a sugar transporter, promoting uptake of glucose into peripheral tissues (Huang and Czech, 2007). Mice with decreased levels of the GLUT4 transporter protein in muscle or adipose tissue show insulin resistance and propensity toward diabetes (Abel et al., 2001; Zisman et al., 2000; Li et al., 2000; Stenbit et al., 1997; Rossetti et al., 1997). Fat cells from transgenic *Drosophila* expressing tagged human GLUT4 respond to insulin by increasing GLUT4 trafficking and translocation to the plasma membrane, showing that *Drosophila* have the necessary signals to direct mammalian-like trafficking in response to insulin signalling (Crivat et al., 2013).

Glucagon action

On the other hand, when sugar levels are low, mammalian pancreatic α cells secrete glucagon and *Drosophila* AKH-producing cells (APCs) secrete the structurally similar AKH (Van der Horst, 2003). Glucagon/AKH bind to their receptors on the liver/fat body, leading to activation of glycogen phosphorylase and consequent release of glucose into circulation (Arrese and Soulages, 2010; Jiang and Zhang, 2003). Knockout of the glucagon receptor in mice or *Akh* in flies results in lower glucose levels, whereas ectopic expression of *Akh* in the fly fat body results in release of stored carbohydrate (Gelling et al., 2003; Van der Horst, 2003; Lee and Park, 2004)

1.10.4 Conserved metabolic pathology: the effect of high-calorie diets

As discussed earlier, the modern obesity epidemic is partially driven by high calorie diets. *Drosophila* develop the same metabolic phenotypes as mammals when given diets high in fat and/or sugar. This includes weight gain, increased fat storage, hyperglycaemia, insulin-resistance, cardiopathy, and decreased life span, suggesting that there is conservation of the underlying pathological molecular mechanisms (Birse et al., 2010; Heinrichsen and Haddad, 2012; Morris et al., 2012a; Musselman et al., 2011; Na et al., 2013).

1.11 Neuronal control of feeding behaviour in *Drosophila*

As discussed in section 1.4, the human brain is key for controlling energy homeostasis and feeding behaviour. The *Drosophila* brain contains approximately 200,000 neurons (Moloney et al., 2010) which is significantly smaller than a human brain which contains approximately 86 billion neurons (Azevedo et al., 2009). However, many parallels exist between the CNS of *Drosophila* and that of humans. Like in humans, the fly brain is composed of specific substructures specialised for function such as sensory processing (optic and antennal lobes), memory (mushroom bodies), and motor output (central complex) (Moloney et al., 2010). Individual *Drosophila* neurons are also similar to human neurons in terms of their physical structure and biochemical signatures (Moloney et al., 2010). Many homologous features are governed by homologous genes in both species including neuron identity and pruning, axon insulation and pathfinding, and neurotransmitters (Held, 2017). These functional and structural similarities of the CNS have been successfully exploited in studies of neurodegenerative disease but have only recently been applied to the study of obesity. Like in humans, the *Drosophila* brain has a key role in regulation of feeding behaviour and many of the genes involved in mammals have homologues in flies. A few examples will now be described.

NPF and food seeking

In humans, NPY plays a key role in the control of energy homeostasis (see section 1.4). NPY is expressed in hypothalamic neurons and its five GPCR receptors (Y1, Y2, Y4, Y5 and Y6) are distributed across the central and peripheral nervous systems (Tatemoto et al., 1982). In *Drosophila*, neuropeptide F (NPF) and short (s)NPF, and their receptors NPFR and sNPFR, show sequence and structural similarity to NPY and the Y receptor family, and are expressed in the central and peripheral nervous systems (Brown et al., 1999; Garczynski et al., 2002; Lee et al., 2004; Mertens et al., 2002).

These NPY/(s)NPF neurons are activated by energy deficit - NPY neurons by ghrelin and NPFR neurons by starvation (Kohno et al., 2003; Wu et al., 2005a,b). Conversely, both NPY and NPF neurons are negatively regulated by signals of energy surplus such as leptin / upd1 and insulin / dILPs (Sahu, 1998; Schwartz et al., 1992; Wu et al., 2005a,b; Beshel et al., 2017).

Increases in NPY/(s)NPF expression has an orexigenic effect. Intracerebroventricular (ICV) injection of NPY or Y1 agonists into mice causes hyperphagia and weight gain (Henry

et al., 2005; Mullins et al., 2001; Sainsbury et al., 1997). NPF and sNPF mainly act on foraging behaviour, mediated by attraction towards food odours. Over-expression of *sNPF* in the *Drosophila* nervous system promotes food intake and causes flies to behave in a manner that mimics starvation (Lee et al., 2004; Root et al., 2011). Expression of *NPF* in wandering larvae causes them to return to continuous feeding, and NPFR activation increases the attractiveness of both food- and aversive-odours and promotes feeding on noxious and unattractive food, partially mediated by ILP-receptor signalling (Wu et al., 2003, 2005b,a; Beshel and Zhong, 2013; Kohno et al., 2003).

Conversely, ablation of NPY neurons in adult mice leads to anorexia (Wu and Palmiter, 2011) and genetic deletion of Y1 in ob/ob mice reduces their food intake (Pralong et al., 2002). Genetic perturbation of NPF and sNPF neurons also leads to changes in food intake, wandering behaviour, and food-search behaviour in response to starvation (Beshel and Zhong, 2013; Lee et al., 2004; Root et al., 2011; Wu et al., 2003).

upd1 and feeding motivation

As discussed in section 1.3, leptin hormone is released by adipose tissue in relation to fat storage levels in humans. It has an anorexigenic effect, in part brought about by inhibiting NPY neurons (Stephens et al., 1995). The *Drosophila* gene *upd1* shows sequence and structural homology to leptin, and binds to the *domeless* receptor found on NPF neurons, leading to suppressed NPF activity (Beshel, 2016; Boulay et al., 2003; Harrison et al., 1998; Wright et al., 2011).

Leptin-deficient rodents and humans are heavier with larger fat stores due to increased food intake but still show increased food-seeking behaviour (Farooqi and O'Rahilly, 2014). All of these phenotypes are also seen in flies with knock-down of *upd1* (Beshel, 2016). Interestingly, this *Drosophila* phenotype could be rescued by expression of the human leptin gene, highlighting the conservation between humans and *Drosophila* (Beshel, 2016).

hugin and meal initiation

Another component of the neural circuitry of feeding is mammalian neuromedin U (NmU) and the *Drosophila* homologue *hugin*. NmU and hugin have similar structures, and are both cleaved into 8mer peptides with similar sequences that bind to homologous GPCRs (Austin et al., 1995; Brighton et al., 2004; Meng et al., 2002; Rosenkilde et al., 2003). Rat NmU is

specifically expressed in the VMN and *hugin* in the subesophageal ganglion, regions involved in regulating feeding (Howard et al., 2000; Melcher and Pankratz, 2005).

Expression of both NmU and *hugin* is down-regulated by fasting (Itskov and Ribeiro, 2013). Accordingly, NmU knockout mice are obese with increased feeding (Bechtold et al., 2009; Hanada et al., 2004), and blocking synaptic activity of *hugin*-expressing neurons using tetanus toxin decreases the time taken to initiate feeding in *Drosophila* (Melcher and Pankratz, 2005; Melcher et al., 2006, 2007). Conversely, ICV administration of NmU, or overexpression of NmU or *hugin*, suppresses feeding (Bechtold et al., 2009; Howard et al., 2000; Kowalski et al., 2005; Melcher and Pankratz, 2005; Melcher et al., 2006, 2007; Schoofs et al., 2014).

Leucokinin and feeding cessation

The tachykinins are a family of small peptides which, along with their receptors, are expressed in the mammalian hypothalamus and have a role in feeding behaviour (Cvetkovic et al., 2003; Achapu et al., 1992; Sahu et al., 1988; Itskov and Ribeiro, 2013). For example, injection of NPK in rats delays feeding initiation (Sahu et al., 1988).

The *Drosophila* protein leucokinin (Lk) shows sequence homology to the tachykinins, and is expressed in the CNS (Radford et al., 2002). According to published data, flies with mutations of *Lk* or its receptor (Lkr) show increased food intake each meal, but compensate by increasing the interval between their meals (Al-Anzi, 2010; Zandawala et al., 2018). The effects of *Lk* on meal termination are independent of hugin and NPF neurons (Al-Anzi, 2010).

Serotinergic neurons and satiety

Serotonin is a neurotransmitter synthesised within the CNS exclusively in the raphe nuclei in the brainstem of mammals and in the R50H05 region of the central brain in *Drosophila* (Albin et al., 2015; Donovan and Tecott, 2013). In both cases, these small groups of neurons project extensively throughout the CNS. One of the targets of the serotonin neurons in *Drosophila* is Lk neurons (Liu et al., 2015).

In rodents, CNS serotonin signalling suppresses feeding. Hyperphagia can be induced by lesioning the raphe nuclei or ICV injection of pharmaceutical agents that block serotonin signalling (Breisch et al., 1976; Geyer et al., 1976; Klitenick and Wirtshafter, 1988; Saller and Stricker, 1976). Conversely, hypophagia results after central injection of serotonin or

its precursor 5HTP, or by blocking serotonin re-uptake (Blundell et al., 1975; Blundell and Latham, 1979; Simansky, 1996)

In *Drosophila*, disruption of serotonin signalling also strongly affects feeding, but in the opposite direction. Acute activation of the R50H05 serotonergic neurons causes sated flies to feed the same way as starved flies (Albin et al., 2015) whereas using a serotonin receptor antagonist to blocking signalling inhibits feeding in larvae (Gasque et al., 2013).

DSK and satiety

DSK is the *Drosophila* homologue of mammalian CCK (Söderberg et al., 2012). CCK is found in both the brain and the gastrointestinal tract and injection into mammals inhibits food intake by producing satiety (Moran, 2000). In the *Drosophila* brain, DSK is produced by IPCs and knockdown of *DSK* in IPCs, or inactivation of IPCs, decreases satiety signalling in flies and hence increases food intake, even when the food is less palatable or bitter (Söderberg et al., 2012).

The *TFAP-2* and *twz* genes regulate expression of DSK (Williams et al., 2014). The human homologues of these two genes do not yet have a proven role in energy homeostasis but they have been associated with BMI by GWAS (Speliotes et al., 2010).

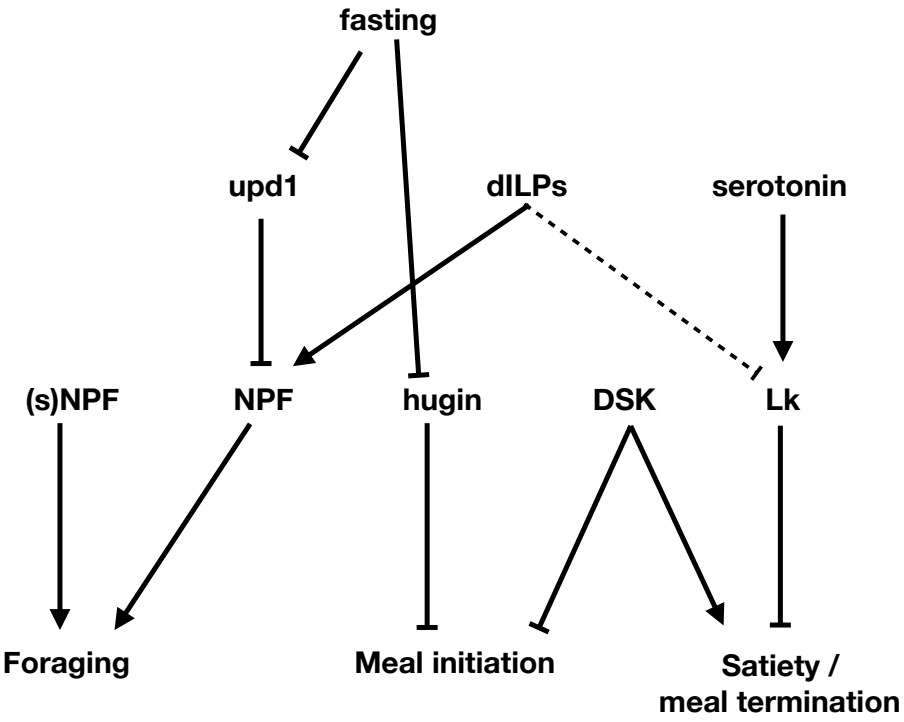


Fig. 1.5 Neural pathways that regulate feeding behaviour in the *Drosophila* adult brain. Figure based on Pool and Scott (2014).
dILP = *Drosophila* insulin-like peptide, DSK = drosulfakinin, Lk = leucokinin, NPF = neuropeptide F, upd1 = unpaired 1.

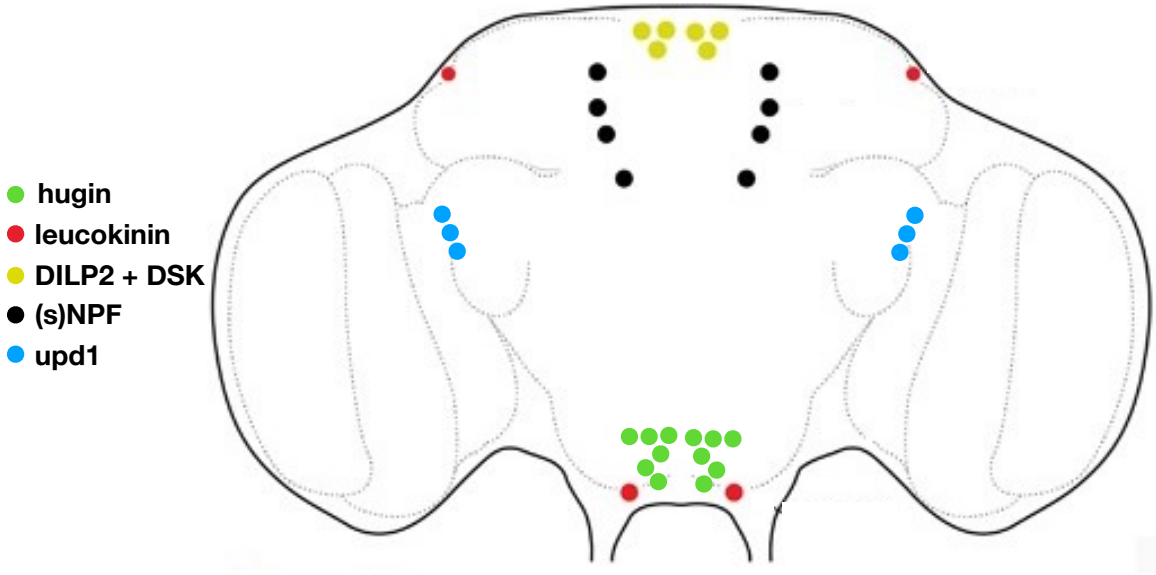


Fig. 1.6 Expression of some neuropeptides involved in feeding behaviour in the *Drosophila* adult brain. Figure shows a coronal section of the *Drosophila* adult brain, adapted from Nässel (2018).
dILP2 = *Drosophila* insulin-like peptide 2, DSK = drosulfakinin, NPF = neuropeptide F, upd1 = unpaired 1.

1.12 High-throughput genetic screens for energy homeostasis phenotypes

1.12.1 Screens in mice

The International Mouse Phenotyping Consortium (IMPC) is currently undertaking a project to produce and phenotype knock-out mouse lines for 20,000 genes (Moore, 2010; Dickinson et al., 2016). The project started in 2011, and has now created mutant mouse lines for 5505 genes. These lines are only partially phenotyped, for example only 594 lines have had food intake measured (data release 8.0, July 2018). Further, because the IMPC are creating whole-organism gene knock-out mutations, there is no phenotype information specific to the brain, and 35% of the lines created so far are lethal before weaning so can not be examined.

Although this data will, one day, be available for all of the genes, the speed of results generation in mice is exponentially slower than that for *Drosophila*. Therefore, at least for the time being, there is need for a high-throughput screen for the role of genes in the brain's control of feeding behaviour and energy homeostasis.

1.12.2 Screens in *Drosophila*

There are many published examples of using *Drosophila* to study one or a few genes involved in energy homeostasis, and some of these were mentioned previously. But as was highlighted in section 1.8, studies using *Drosophila* have the potential to be high-throughput. This, coupled with the well-developed genetic techniques available in flies and the high level of genetic and metabolic conservation means that *Drosophila* have also recently been used for large-scale genetic screens of metabolic phenotypes.

As just one example, Pospisilik et al. (2010) performed a genome-wide screen for TAG levels in adult flies using RNAi knockdown. Overall, the top hit was the *hedgehog* signalling pathway in the fat body. They went on to show that mice with activation of the same pathway in adipose tissue show a near total loss of white fat due to blocked differentiation of white adipocytes. Such studies are "fishing" experiments in which there is no starting hypothesis and results from *Drosophila* physiology are then used to look for results in mice/humans. Although this is often successful (see also screens for metabolic phenotypes by Reis et al. (2010) and Baumbach et al. (2014)), there is no guarantee that results will be applicable to human disease. A different approach was used in this thesis: human genetics are already known but require an animal model for functional investigation.

A PubMed search revealed two published large-scale (>50 genes) *Drosophila*-based screens which started with genes from human metabolic disease. Pendse et al. (2013) and Baranski et al. (2018) studied genetic loci that have been associated by GWAS with type 2 diabetes and BMI respectively in humans. These studies will be explored in more detail in chapter 4, but briefly, Pendse et al. looked for a sucrose-dependent toxicity phenotype, and Baranski et al. for a TAG phenotype. Based on their results, both were able to suggest which genes should be the focus of further studies in humans/mice and in many cases these were not the genes which would have been selected based solely on the original human statistics-based data, highlighting the utility of *Drosophila* for unbiased screening of genes. However, these studies assessed just one phenotype and did not look at feeding behaviour.

In terms of feeding behaviour, a PubMed search found only two published large-scale screens in *Drosophila*. The first screen used a collection of lines (not genome-wide) to look for larval feeding defects and identified the genes *ppl* and *klu* (Melcher and Pankratz, 2005; Zinke et al., 1999). The second screen used *Drosophila* larvae to screen 3630 drugs for inhibition of food intake, leading to identification of an anorectic serotonin receptor antagonist called metitepine (Gasque et al., 2013). As above, these screens may or may not provide information applicable to human disease.

Thus, to my knowledge, there are no published large-scale screens of human genes using *Drosophila* to look for a role in the neuronal control of energy homeostasis and in particular feeding behaviour.

1.13 Project aims

The public health concerns over obesity combined with the lack of a full understanding of the mechanisms of energy homeostasis are the primary motivation of this work. This thesis aimed to develop an *in vivo*, high-throughput, *Drosophila*-based screen to assess large lists of human genes for a role in the neuronal control of energy homeostasis and in particular feeding behaviour, and then to use the screen to identify top candidates for further studies in mice and humans. The aims of each chapter are to:

- Chapter 3: Develop and validate a *Drosophila*-based screen for looking at the *in vivo* impact of genes on energy homeostasis and feeding behaviour using readily available *Drosophila* lines
- Chapter 4: use this high-throughput screen to analyse genes suggested by GWAS to affect BMI. This list of candidate genes has already been probed but only to a limited extent, making it ideal for use as both proof-of-principle as well as a potential source of novel results.
- Chapter 5: Explore the possibility of creating specific *de novo* genetic mutations in *Drosophila* for use in the screen
- Chapter 6: Use the final screen developed in chapters 3 through 5 to explore genes from the hypothalamus which have been shown to be regulated by fasting, and thus identify novel genes for further investigation.

Chapter 2

Materials and Methods

2.1 Fly husbandry

The different diet media to which flies were exposed are listed in the table below. All flies were raised on ND. Adult flies for phenotyping were kept on either ND, HSD, HFD, starvation media, or dye food, as appropriate.

Diet	Name	Ingredients
Normal diet	ND	agar, dextrose, maize, yeast, nipagin, water
High fat diet	HFD	ND + 20% coconut oil
High sucrose diet	HSD	ND + 35% sucrose
Starvation media	SM	agar and water only
Dye food	-	ND + 1% Fast Green FCF dye (Sigma)

All experimental flies were kept at 25^{o}C and 60-70% humidity with a 14 hour light / 10 hour dark cycle, except for flies on HFD which were kept at 20^{o}C, variable humidity, in the dark.

2.2 Lines used

All lines used in this thesis are listed in table 2.1. Lines were purchased from VDRC, Bloomington, Kyoto, Harvard, or provided by the University of Cambridge fly facility.

Label	*Drosophila* gene	CGNR	Stock centre ID	Human gene
Vienna Drosophila Resource Center (VDRC, www.vdrc.at)				
V1	*CG10920*	CG10920	109327KK	MTCH2
V3	*fuss*	CG11093	103367KK	LBXCOR1
V4	*St2*	CG16733	47019GD	SULT1A2
V5	*PKN*	CG11221	100163KK	SBK1
V6/VV59	*CG4945*	CG4945	24683GD	SBK1
V7	*CG8916*	CG8916	101633KK	GABRG1
V8	*CG6550*	CG6550	4947GD	CDKAL1
V9	*dSp1*	CG1343	12607GD	KLF9
V10	*cabot*	CG4427	15555GD	KLF9
V11	*rab21*	CG17515	109991KK	RBJ
V12	*mIF3*	CG13163	39785GD	MTIF3
V13	*boss*	CG8285	4365GD	GPRC5B
V14	*Ets96B*	CG6892	30552GD	ETV5
V15	*dSec16*	CG32654	29635GD	SEC16B
V16	*NT1*	CG42576	108894KK	BDNF
VV1	*psn*	CG18803	101379KK	PSEN1
VV2	*engrailed*	CG9015	105678KK	EN1
VV3	*upd1*	CG5993	3282GD	LEPTIN
VV5	*fz*	CG17697	43075GD	FZD3
VV6	*Ptp61F*	CG9181	37436GD	PTPN2
VV7	*trn*	CG11280	5242GD	LRRN6C
VV8	*HGTX*	CG13475	12608GD	NKX6.1
VV9	*fur1*	CG10772	22853GD	PCSK1
VV10	*caps*	CG11282	27097GD	LRRN6C
VV11	*cert*	CG7207	27914GD	COL4A3BP
VV12	*shtd*	CG9198	29072GD	ANAPC1
VV13	*Fancl*	CG12812	32025GD	FANCL
VV14	*Ac3*	CG1506	33217GD	ADCY3
VV15	*heph*	CG31000	33735GD	PTBP2
VV16	*rab2*	CG3269	34767GD	PCSK1
VV17	*Lrp1*	CG33087	39215GD	LRP1B
VV18	*dsor1*	CG15793	40026GD	KCTD15
VV19	*TfAP-2*	CG7807	41130GD	TFAP2B

Label	***Drosophila* gene**	**CGNR**	**Stock centre ID**	**Human gene**
VV20/VV57	*sbb*	CG5580	41845GD	ZNF608
VV21	*Tmem18*	CG30051	43067GD	TMEM18
VV22	*Fili*	CG34368	44532GD	LRRN6C
VV23	*RpL27A*	CG15442	46439GD	RPL27A
VV24	*mEFTu1*	CG6050	48981GD	TUFM
VV25	*isoQC*	CG5976	101533KK	QPCTL
VV26	*CG3732*	CG3732	105968KK	ZRANB2
VV27	*Oscillin*	CG6957	106685KK	GNPDA2
VV28	*Nmda1*	CG3798	108378KK	FAIM2
VV29	*Hmgcr*	CG10367	108617KK	HMGCR
VV30	*Atx2*	CG5166	108843KK	ATXN2L
VV31	*RtGEF*	CG10043	17966GD	SH3YL1
VV32	*hth*	CG17117	100630KK	MEIS1
VV33	*DSK*	CG18090	14201GD	CCK
VV34	*unc13*	CG2999	33606GD	UNC13A
VV35	*unc13-4a*	CG32381	109304KK	UNC13A
VV36	*DSK*	CG18090	106592KK	CCK
VV37	*CG2118*	CG2118	25406GD	MCCC1
VV40	*sns*	CG33141	109442KK	CADM2
VV41	*Dh44-R1*	CG8422	110708KK	GIPR
VV42	*Dh44-R2*	CG12370	43313GD	GIPR
VV43	*Mtch*	CG6851	44305GD	MTCH2
VV44	*Lac*	CG12369	35524GD	NEGR1
VV45	*DIP-iota*	CG11320	18054GD	NEGR1
VV46	*Ama*	CG2198	22944GD	NEGR1
VV47	*Nrx-1*	CG7050	4306GD	NRXN3
VV48	*Aps*	CG6391	110607KK	NUDT3
VV49	*Fur2*	CG18734	1021GD	PCSK1
VV50	*QC*	CG32412	38277GD	QPCTL
VV51	*Zip71B*	CG10006	44538GD	SLC39A8
VV52	*Syn2*	CG4905	110602KK	SNTG2
VV53	*TfAP-2*	CG7807	41130GD	TFAP2B
VV54	*twz*	CG10440	110265KK	KCTD15
VV55	*Grd*	CG7446	5329GD	GABRG1
VV56	*CG14762*	CG14762	31014GD	LRRN6C

Label	***Drosophila* gene**	**CGNR**	**Stock centre ID**	**Human gene**
VV58	*PKD*	CG7125	106255KK	PRKD1
VV60	*mIF3*	CG13163	107373KK	MTIF3
VV70	*prosap*	CG30483	44830GD	SHANK3
VV71	*auxilin*	CG1107	16182GD	GAK
VV72	*chico*	CG5686	7776GD	IRS1
VV74	*hugin*	CG6371	26766GD	Neuromedin U
VV75	*mRpS21*	CG32854	101430KK	MRPS21
VV78	*Ect3*	CG3132	16779GD	GLB1
VV79	*KDM4A*	CG15835	107868KK	KDM4C
VV80	*poe*	CG14472	17648GD	UBR4
VT1	*betaTub56D*	CG9277	109736KK	Tubb4b
VT2	*Arpc5*	CG9881	28141GD	Arpc5l
VT3	*spidey*	CG1444	40949GD	Hsd17b12
VT4	*Rep*	CG8432	28866GD	Chml
VT5	*CG1764*	CG1764	20507GD	Ddah1
VT6	*CG9231*	CG9231	9101GD	Fam162a
VT7	*Rgk1*	CG44011	25574GD	Gem
VT8	*SIFaR*	CG10823	1783GD	Hcrtr2
VT9	*GCS2β*	CG6453	37991GD	Prkcsh
VT10	*CG9267*	CG9267	2879GD	Ptplad1
VT11	*Spred*	CG10155	18025GD	Spred2 / Spred3
VT12	*Tapδ*	CG9035	8759GD	Ssr4
VT14	*put*	CG7904	848GD	Acvr2a
VT15	*PMCA*	CG42314	30203GD	Atp2b2
VT16	*yrt*	CG9764	28674GD	Epb4.1l4b
VT17	*CG2656*	CG2656	25423GD	Gpn3
VT18	*na*	CG1517	3307GD	Nalcn
VT19	*Nmt*	CG7436	28019GD	Nmt1
VT20	*O-fut1*	CG12366	44045GD	Pofut1
VT21	*AMPKα*	CG3051	1827GD	Prkaa1
VT22	*Purα*	CG1507	12765GD	Pura
VT23	*GlyP*	CG7254	27928GD	Pygl
VT24	*Syt7*	CG2381	24988GD	Syt7
VT25	*ktub*	CG9398	29110GD	Tub
VT26	*Vps50*	CG4996	34913GD	Ccdc132

Label	***Drosophila* gene**	**CGNR**	**Stock centre ID**	**Human gene**
VT27	*kug*	CG7749	3749GD	Fat3
VT28	*Irk2*	CG4370	4341GD	Kcnj3
VT29	*KCNQ*	CG33135	8754GD	Kcnq3
VT30	*CG18301*	CG18301	31023GD	Lipa
VT31	*Ctl2*	CG11880	22867GD	Slc44a5
VT32	*CG32052*	CG32052	21437GD	Smpdl3b
VT33	*btsz*	CG44012	35205GD	Sytl4
VT34	*Vps39*	CG7146	40425GD	Vps39
VT35	*CG17896*	CG17896	5581GD	Aldh6a1
VT36	*Eip63E*	CG10579	47860GD	Cdk14
VT37	*CG31064*	CG31064	33754GD	Rufy2
VT38	*Caβ*	CG42403	102188KK	Cacnb4
VT39	*DhpD*	CG18143	106096KK	Gda
VT40	*Nedd4*	CG42279	108475KK	Nedd4l
VT41	*S6kII*	CG17596	101451KK	Rps6ka6
VT42	*Gat*	CG1732	106638KK	Slc6a1
VT43	*Lrch*	CG6860	107047KK	Lrch2
VT44	*Eip75B*	CG8127	108399KK	Nr1d2
VT45	*JIL-1*	CG6297	107001KK	Rps6ka5
VT46	*Dyb*	CG8529	104485KK	Dtna
VT47	*Hs6st*	CG4451	110424KK	Hs6st2
VT48	*Rbfox1*	CG32062	110518KK	Rbfox1
VT49	*CG7369*	CG7369	100824KK	Rasgef1a / Rasgef1c
VT50	*CG4853*	CG4853	19877GD	Rasgef1a / Rasgef1c
VT51	*5-HT1A*	CG16720	106094KK	Htr1a
VT52	*5-HT1B*	CG15113	109929KK	Htr1a
VT53	*Rbp9*	CG3151	101412KK	Elavl2
VT54	*fne*	CG4396	101508KK	Elavl2
KK	n/a	n/a	60100	n/a
GD	n/a	n/a	60000	n/a
VV4	KK_{tiptop}	n/a	60101	n/a
KYOTO Stock Center (DGRC, www.kyotofly.kit.jp)				
K1	*Aps*	CG6391	103540	NUDT3
K2	*Ama*	CG2198	103970	NEGR1
K3	*fur2*	CG18734	104593	PCSK1

Label	*Drosophila* gene	CGNR	Stock centre ID	Human gene
K4	*Fancl*	CG12812	140133	FANCL
K5	*Tmem18*	CG30051	141615	TMEM18
K6	*Zip71B*	CG10006	207317	SLC39A8
KK1	*unc13*	CG2999	101911	UNC13A
KK2	*psn*	CG18803	123495	PSEN1
Exelixis Collection at the Harvard Medical School (www.drosophila.med.harvard.edu)				
H1	*isoQC*	CG5976	c04389	QPCTL
H2	*CG3732*	CG3732	c06521	ZRANB2
H3	*CG11342*	CG11342	f06786	BCDIN3D
H4	*cert*	CG7207	f04650	COL4A3BP
H5	*Oscillin*	CG6957	e01629	GNPDA2
HH1	*DSK*	CG18090	f02648	CCK
HH2	*hip14*	CG6017	d07562	Hip14
Bloomington Drosophila stock centre (www.bdsc.indiana.edu)				
B1	*heph*	CG31000	635	PTBP2
B2	*trn*	CG11280	4550	LRRN6C
B3	*dsor1*	CG7693	5545	MAP2K5
B4	*Rpl27A*	CG15442	5697	RPL27A
B7	*shtd*	CG9198	9243	APC1
B8	*Hmgcr*	CG10367	11522	HMGCR
B9	*caps*	CG11282	11579	LRRN6C
B10	*mEFTu1*	CG6050	12215	TUFM
B11	*sbb*	CG5580	12772	ZNF608
B12	*Lac*	CG12369	14577	NEGR1
B13	*Lrp1*	CG33087	16864	LRP1B
B14	*HGTX*	CG13475	19133	NKX6.1
B15	*rab2*	CG3269	19993	PCSK1
B16	*Atx2*	CG5166	21645	ATXN2L
B17	*Nrx-1*	CG7050	21977	NRXN3
B18	*TfAP-2*	CG7807	22901	TFAP2B
B19	*DIP-iota*	CG11320	23405	NEGR1
B20	*Dh44-R1*	CG8422	23517	GIPR
B21	*twz*	CG10440	25846	KCTD15
B22	*Mtch*	CG6851	27981	MTCH2
B23	*QC*	CG32412	29127	QPCTL

Label	***Drosophila* gene**	**CGNR**	**Stock centre ID**	**Human gene**
B24	*Grd*	CG7446	30596	GABRG1
B25	*sns*	CG33141	35916	CADM2
B26	*TfAP-2*	CG7807	36132	TFAP2B
B27	*Syn2*	CG4905	37052	SNTG2
B28	*fili*	CG34368	38026	LRRN6C
B29	*CG14762*	CG14762	44774	LRRN6C
B30	*PKD*	CG7125	52147	PRKD1
B31	*Dh44-R2*	CG12370	53102	GIPR
B32	*NMDA1*	CG3798	53122	FAIM2
B33	*fur1*	CG10772	53406	PCSK1
B34	*RtGEF*	CG10043	53466	SH3YL1
B35	*Ac3*	CG1506	53472	ADCY3
BB1	*engrailed*	CG9015	265	EN1
BB2	*frizzled*	CG17697	1676	FZD3
BB3	*engrailed*	CG9015	1817	EN1
BB4	*upd1*	CG5993	4767	Leptin
BB5	*psn*	CG18803	5463	PSEN1
BB6	*psn*	CG18803	8299	PSEN1
BB7	*chico*	CG5686	10738	IRS1
BB8	*chico*	CG5686	14337	IRS1
BB9	*Btbd9*	CG1826	14666	BTBD9
BB10	*bmm*	CG5295	15959	ATGL
BB11	*leucokinin*	CG13480	16324	Tachykinin
BB12	*prosap*	CG30483	17047	SHANK3
BB13	*DSK*	CG18090	17146	CCK
BB14	*Hip14*	CG6017	17614	Hip14
BB15	*nACHRα5*	CG8178	18381	CHRNA7
BB16	*mfrn*	CG4963	19811	SLC25A37
BB17	*hugin*	CG6371	23491	Neuromedin U
BB18	*prosap*	CG30483	24446	SHANK3
BB19	*auxilin*	CG1107	25674	GAK
BB20	*lap*	CG2520	26463	PICALM
BB21	*auxilin*	CG1107	30174	GAK
BB22	*hugin*	CG6371	34419	Neuromedin U
BB23	*upd2*	CG5988	55727	Leptin

Label	***Drosophila* gene**	**CGNR**	**Stock centre ID**	**Human gene**
BB24	*bmm*	CG5295	58659	ATGL
BB25	*nACHRα5*	CG8178	59218	CHRNA7
BB26	*mfrn*	CG4963	59729	SLC25A37
BB27	*bmm*	CG5295	62615	ATGL
BB28	*tau*	CG45110	63271	MAPT
BB29	*tau*	CG45110	64782	MAPT
BB30	*tht*	CG17117	65540	MEIS1

University of Cambridge fly facility (www.flyfacility.gen.cam.ac.uk)

Label	**Genotype**
B5	*elav-GAL4/elav-GAL4 (III)*
B6	*elav-GAL4/CyO*
ACT	*Act5C-GAL4/CyO*
DB	*if / CyO ; TM6b / MKRS*
CFD1	*y1 P(act5c-cas9, w+) M(3xP3-RFP.attP)ZH-2A w*(I)*
FM7	*FM7 / w^a*
w^{1118}	*w^{1118}*
yw	*$y^1 w^1$*
JP084	*UAS_G-Cas9 / TM6b*
nanos-Cas9	*y[1] sc[1] v[1]; (y[+t7.7] v[+t1.8]=nanos-Cas9)attp40*
PhiC31+attP	*y w M(eGFP, vas-int, dmRFP)ZH-2A; P(CaryP)attP40*

2.3 Experimental flies

2.3.1 Homologue identification

Drosophila homologues of the human / mouse genes of interest were identified using the ENSEMBL orthologue tool (Zerbino et al., 2018), and FlyBase BLAST searches of protein sequences derived from NCBI (Gramates et al., 2017).

2.3.2 Generation of experimental RNAi flies

UAS_G-siRNA lines were crossed to GAL4 lines (elav-GAL4 or act-GAL4). Controls were created by crossing each GAL4 line individually to KK, KK_{tiptop} and GD. Experimental flies were compared to the relevant control background. To standardise the effects of parental environment on offspring fitness, all UAS_G-siRNA stocks were kept in bottles at an approximately constant density. To standardise the effects of parental age on offspring fitness, crosses were set up using flies which were 1-5 days old.

On day 1, 5 UAS_G-siRNA female virgins were placed in a vial of ND at 25^oC with 2 elav-GAL4 or 5 act-GAL4 males. On day 4 these parents were removed. On day 14 offspring were transferred to a new vial and allowed to mate, as mating alters gene expression and metabolic parameters (Ellis and Carney, 2010; Papanastasiou et al., 2013). On day 15 the females were removed, and on day 16 male flies were placed on SM for 3 hours to synchronise their metabolism, before being returned to the appropriate diet for the experiment.

2.3.3 Generation of experimental LoF flies

Stocks were kept in bottles at approximately constant density. Bottles were cleared and newly eclosed males collected every 24 hours. Immediately after collection, flies were placed on SM for 3 hours to synchronise their metabolism, before being moved to the appropriate diet for the experiment. w^{1118} and y^1w^1 were used as background control lines.

2.4 Assays to phenotype *Drosophila*

2.4.1 Timeline

During the 3 hour synchronisation starvation, flies were counted into groups for assaying. All assays were done at a certain time relative to the synchronisation (day 0), because *Drosophila* activity and feeding behaviours are affected by age and circadian rhythms (Xu et al., 2008; Iliadi and Boulianne, 2010).

Assay	No of flies / repeat	Assay day	Assay time
qPCR	10	4	10am
Wet mass, TAG, glucose	15	5	10am
Dry mass	15	5 + 8	3.30pm
Starvation resistance	15	4 + 7	10am
Dye feeding over-feeding	15	4 + 5	9am
Dye feeding absoprtion	15	8 + 9	9am
CAFE	8	4 + 5	2pm
Climbing	15	4	3.30pm
Feeding motivation	1	5	10am

2.4.2 Repeats and statistics

For LoF flies, each assay was repeated 7 times using flies from independent collections. For RNAi flies, each assay was repeated 5 times on independent crosses. Data was analysed by unpaired homoscedastic Student t-test, except where stated otherwise. Based on the use of the protocols in the literature, the assays may have standard deviation up to 25%. In this case, sample sizes of 5 repeats per experimental group will be able to detect changes of 50% at a significance level of $p<0.05$ (two- tailed) and 80% power (Kadam and Bhalerao, 2010).

2.4.3 Wet mass

Groups of 15 *Drosophila* were frozen on dry ice and weighed using a microbalance, then stored at -80^{o}C for TAG/glucose analysis.

2.4.4 Dry mass

Flies in groups of 15 were frozen on dry ice then placed in a 95^{o}C heat-block for 6 minutes (1 minute with the lid closed, 5 minutes with it open). Flies were kept in eppendorf tubes at room temperature for 3 days to dessicate and then weighed using a microbalance.

2.4.5 TAG and glucose levels

Frozen *Drosophila* were placed in FastPrep tubes containing Lysis Beads and Matrix D (MP Biomedicals) and 350μl of cold PBST (PBS + 0.05% Tween-20) then homogenised using a FastPrep-24 homogeniser (MP Biomedicals) for 60s at 6m/s. Solutions were centrifuged (16100rcf, 4^{o}C, 3 minutes) to pellet debris and 300μl of supernatant pipetted into a fresh Eppendorf on ice. Homogenates were heat-inactivated (5 minutes, 70^{o}C) then 50μl transferred into a fresh tube and stored at -80^{o}C for TAG analysis. The remaining supernatant was re-centrifuged (16100rcf, 4^{o}C, 3 minutes) to pellet debris and 100μl transferred into a fresh tube and stored at -80^{o}C for glucose analysis. TAG analysis and glucose analysis was done by the Cambridge Core Biochemical Assay Laboratory (www.cuh.nhs.uk/core-biochemical-assay-laboratory) using enzymatic assays. TAG and glucose amounts were normalised to number of *Drosophila*.

2.4.6 Dye food over-feeding

Drosophila were transferred onto: SM for 24 hours, ND for 20 minutes and then dye food for 16 minutes. The number of flies with visible dye in their abdomen (mid-gut and/or crop) was counted as a percentage of the total number.

2.4.7 Dye food absorption

After the first dye food assay, flies were returned to ND for 48 hours then dye food for 24 hours. Anaesthetised flies were frozen on dry ice and stored at -80^{o}C. Samples were transferred into FastPrep tubes containing Lysis Beads and Matrix D and 300μl of PBST (PBS + 1% triton X) then homogenised using a FastPrep-24 homogeniser (30s, 6m/s). Samples were centrifuged (1500g, RT, 3 minutes) and 50μl of supernatant pipetted into each of 2 wells of a CoStar 96-well flat bottom UV-transparent plate (Corning). The absorbance at 625nm was measured using a microplate reader (TECAN Infinite M1000 PRO). Each sample

was repeated twice and an average taken. The amount of dye was normalised to the number of flies.

2.4.8 CAFE assay

Flies were placed onto SM. Specialised caps (made by the Cambridge University Psychology Department Electronics Workshop) were inserted into the vials. Per vial, two 5μl capillary tubes were filled with liquid food (9ml water + 0.5g sucrose + 0.5g yeast extract) via capillary action. The top of the meniscus was marked, and the filled capillaries were inserted into the lid, and then left at 25^{o}C for 24 hours. The movement of the meniscus was measured and evaporation (as measured by a vial containing no flies) subtracted to give the volume of food eaten. This value was normalised to the number of flies.

2.4.9 Climbing assay

Flies were placed in empty vials with a line 4.5cm from the bottom (mid-way). Vials were placed in front of a white background, tapped 3 times to displace flies to the bottom surface, and recorded for 10s using a Sony DCR-SR32 camera. Videos were analysed and the number of flies past each line was counted at 75 frames (3s) after displacement. Flies on the line were counted as above.

2.4.10 Starvation resistance

Drosophila were placed in vials containing SM. The number of dead flies was counted at 4pm every day until all flies were dead. Flies were transferred into new vials of SM every other day, noting how many dead flies transferred into the new tube, and whether any live flies escaped. Vials were kept inverted to prevent flies becoming trapped in the media.

2.4.11 Feeding motivation

At one end of each lane of a 14-well plate, 10μl of a 10^{-6} dilution of benzaldehyde solution (Sigma) was placed on top of a small ball of cotton wool to act as a repellent. A small blob of fresh yeast paste food was then placed in between the repellent and the flies. Individual flies were anaesthetised using ice and placed into each lane. Hardware and software to track the movement of flies was developed by Zantiks (www.zantiks.com). Each lane was split

into 3 equal size zones, and the time that each fly spent within each zone was measured over a 1 hour period. The tracking script is given in figure 2.1.

```
# Zanscript on-line development - enter script below.

INCLUDE zsys



ACTION MAIN

    DEFINE NUMTRIALS 12
    DEFINE MEASURETIME 300

    Set(Tile_size,4)
    Set(AUTOREF_MODE,1)
    Set(AUTOREF_TIMEOUT,1)
    Set(DETECTOR_THRESHOLD,7)
    Set(SEARCH_DISTANCE,10)
    Set(SEARCH_STEP,3)
    Set(FILTER_RADIUS,5)
    Set(maximum_targets,1)


    Load(ARENAS,"arena_14_tung.bmp")
    Load(DETECTORS,"arena14_3zone_tung.bmp")

    LOGCREATE("TEXT:TIME|TEXT:|TEXT:|TEXT:|TEXT:A1 Z1|TEXT:A1 Z2|TEXT:A1 Z3")
    LOGAPPEND("TEXT:B1 Z1|TEXT:B1 Z2|TEXT:B1 Z3")
    LOGAPPEND("TEXT:C1 Z1|TEXT:C1 Z2|TEXT:C1 Z3")
    LOGAPPEND("TEXT:D1 Z1|TEXT:D1 Z2|TEXT:D1 Z3")
    LOGAPPEND("TEXT:E1 Z1|TEXT:E1 Z2|TEXT:E1 Z3")
    LOGAPPEND("TEXT:F1 Z1|TEXT:F1 Z2|TEXT:F1 Z3")
    LOGAPPEND("TEXT:G1 Z1|TEXT:G1 Z2|TEXT:G1 Z3")
    LOGAPPEND("TEXT:H1 Z1|TEXT:H1 Z2|TEXT:H1 Z3")
    LOGAPPEND("TEXT:I1 Z1|TEXT:I1 Z2|TEXT:I1 Z3")
    LOGAPPEND("TEXT:J1 Z1|TEXT:J1 Z2|TEXT:J1 Z3")
    LOGAPPEND("TEXT:K1 Z1|TEXT:K1 Z2|TEXT:K1 Z3")
    LOGAPPEND("TEXT:L1 Z1|TEXT:L1 Z2|TEXT:L1 Z3")
    LOGAPPEND("TEXT:M1 Z1|TEXT:M1 Z2|TEXT:M1 Z3")
    LOGAPPEND("TEXT:N1 Z1|TEXT:N1 Z2|TEXT:N1 Z3")
    LOGRUN()




    AUTOREFERENCE(0)
    INVOKE(MEASURE,NUMTRIALS)

COMPLETE


ACTION MEASURE
    LOGDATA(DATA_SNAPSHOT,"beginsample")
        VIDEO(30,"14wellzones")
    WAIT(MEASURETIME)

    LOGDATA(DATA_SNAPSHOT,"endsample")
    LOGDATA(DATA_SELECT,"beginsample")
    LOGDATA(DATA_DELTA,"endsample")

     LOGCREATE("RUNTIME|TEXT:|TEXT:|TEXT:|TEXT:ZONE_DISTANCES|ZONE_DISTANCES:A1-14 Z1-3")
```

Fig. 2.1 Coding developed in collaboration with Zantiks to measure feeding motivation in adult *Drosophila*.

2.4.12 Larval assays

Cages were set up containing 60 virgin female and 30 male flies on apple juice agar plates with fresh yeast paste and left for 3 days at 25^{o}C. On the day of collection, a new plate with fresh yeast paste was given every 30 minutes for 1 hour to clear old embryos from the oviducts of the female flies. Embryos were collected on a fresh plate for 3 hours and then left for 3 days to develop into 3rd instar larvae. For the assay, larvae were either fasted in distilled water for 2 hours or fed in liquid yeast for 2 hours. Yeast paste (6ml of distilled water + 6g of active dry yeast) was placed in the centre of, and 20 larvae were placed on the inside rim of, an apple juice agar plate. The number of larvae in/out of the food was counted after 20 minutes.

2.5 Gene editing using CRISPR

2.5.1 gRNA design

gRNAs to target the *white* gene were taken from the literature (Ge et al., 2016). For other genes of interest, possible gRNA sites were identified using the online CRISPOR tool (Haeussler et al., 2016) and the final gRNAs chosen based on the following criteria: 20bp long, Fusi-Doench score >50, no predicted off-targets with 3 or fewer mismatches, position near the beginning/end of the gene, presence in all isoforms, avoidance of overlapping genes. The sequences used were:

Gene	gRNA-1 sequence	gRNA-2 sequence
white	ataccattcctgctctttgg	gtccagcggcagatcggcgg
rab21	cagccgcaagatgtccctgg	ttattactgggcgctcagga
fuss	actcttgaccttcaatgtgca	aataagtggagttatccagt

For each target gene (*fuss, rab21, white*), 1 plasmid was created each containing 2 gRNA sequences

2.5.2 Creation of the gRNA-tRNA insert

The gRNA-tRNA insert was created by PCR using the following primers, which contain the gRNAs designed above:

Gene	Primer	Sequence
white	F	gcggcccgggttcgattcccggccgatgcaataccattcctgctctttgggttttagagctagaaatagcaag
	R	attttaacttgctatttctagctctaaaacccgccgatctgccgctggactgcaccagccgggaatcgaaccc
rab21	F	gcggcccgggttcgattcccggccgatgcacagccgcaagatgtccctgggttttagagctagaaatagcaag
	R	attttaacttgctatttctagctctaaaactcctgagcgcccagtaataatgcaccagccgggaatcgaaccc
fuss	F	gcggcccgggttcgattcccggccgatgcactcttgaccttcaatgtgcagttttagagctagaaatagcaag
	R	attttaacttgctatttctagctctaaaacactggataactccacttatttgcaccagccgggaatcgaaccc

Amplification of the insert was done using pCFD5 as a template. 25μl of Q5 Hot Start High-Fidelity 2x Master Mix (NEB), 2.5μl of 10μM primer F, 2.5μl of 10μM primer R, 1ng of pCFD5 and to 50μl of nuclease-free water were mixed by pipetting. The reaction was incubated as follows: 98oC fo 30s, then 32 cycles of 98^{o}C for 10s, 61^{o}C (increasing by 0.5^{o}C each cycle up to 72^{o}C) and 72^{o}C for 20s, then 72^{o}C for 120s. PCR products were separated by agarose gel electrophoresis and bands of the correct size (234bp) were excised

and purified using a Qiaquick gel extraction kit (Qiagen) according to the manufacturer's instructions.

2.5.3 Plasmid digestion

Plasmid pCFD5 was a gift from Simon Bullock (Addgene plasmid 73914). pCFD5-containing cultures were grown in 2ml of LB broth + 100μg/ml Ampicillin overnight in a 37^{o}C shaking incubator. pCFD5 was extracted using QIAprep Miniprep kit (Qiagen) according to the manufacturer's instructions, and then digested using Bpil (Thermo Scientific). Digested plasmid was separated from undigested plasmid by electrophoresis (80V, 1.5 hours, 1% agarose gel), cut out of the gel under UV light, and then extracted using a Qiaquick gel extraction kit (Qiagen) according to the manufacturer's instructions.

2.5.4 Insertion of gRNA-tRNA into pCFD5

Digested pCFD5 plasmid was combined with the PCR inserts by a Gibson Assembly reaction using Gibson Assembly Master Mix 2x (NEB) according to the manufacturer's instructions. Plasmids were transformed by heat shock into 5-alpha Competent *E. coli* cells (NEB). The cells were streaked onto LB$^{\text{ampicillin}}$ plates and grown overnight at 37^{o}C. Individual colonies were cultured in 2ml of LB broth + 100μg/ml Ampicillin overnight in a 37^{o}C shaking incubator. The resulting cultures were purified using a GeneJET Plasmid Miniprep Kit (Thermo Scientific) according to manufacturer's instructions.

Production of the expected plasmid was confirmed by 2 sequencing reaction using either primer pCFD5 F (cgactgggtaaatggtcctg) or primer pCFD5 R (ggcgaacagagatgagatttgac), and Big Dye Terminator v3.1 Cycle (Thermo Fischer) according to the manufacturer's instructions. Reaction mixtures were cleaned using AgenCourt CleanSEQ magnetic beads according to the manufacturer's instructions, and then sequenced using a ABI 3730 DNA Sequencer (48 Cap). Sequences were analysed using ApE software.

Successful plasmids were re-transformed and cultured, and then purified using a HiSpeed Plasmid Midi kit (Qiagen) according to the manufacturer's instructions.

2.5.5 Embryo injection

gRNA-tRNA plasmids were micro-injected into dechorionated syncytial embryos by the Cambridge University *Drosophila* microinjection service, according to the following table:

Plasmids injected	Concentration	*Drosophila* line injected
fuss + white	100ng/μl	nanos-Cas9
rab21 + white	100ng/μl	nanos-Cas9
fuss	400ng/μl	PhiC31+attP
rab21	400ng/μl	PhiC31+attP

Embryos were grown to adulthood, crossed to balancers, and the offspring genotyped by eye colour, PCR, and sequencing.

2.6 Genotyping of *Drosophila* by PCR

2.6.1 DNA extraction

Individual flies were placed into a PCR tube and mashed using a pipette tip with 50μl of sterile Squish Buffer (9.8ml of nuclease-free water, 100μl of 1M Tris pH 8.0, 20μl 0.5M of EDTA, 50μl 5M of NaCl, and 0.1μl of 10mg/ml Proteinase K (Sigma)). The solution was incubated at 37^{o}C for 1 hour to release the DNA and then heated to 95^{o}C for 3 minutes to inactivate the Proteinase K. 50μl of solution was transferred to a new tube and diluted with 50μl of nuclease-free water.

2.6.2 PCR reaction

Per sample, 25μl of 2x Terra PCR Direct Buffer (Takara), 2μl of 10μM primer F, 2μl of 10μM primer R, 5μl of DNA extract (see above), 1μl of Terra PCR direct polymerase mix (Takara), and 16μl of nuclease-free water were mixed gently by pipetting. The primer sequences were:

Gene	Primer F sequence	Primer R sequence
fuss	cgactgggtaaatggtcctg	ggcgaacagagatgagatttgac
rab21	gcgtatcaaggtaggacacc	atcacaatgggaacggctaa

For *rab21*, a 3-step PCR was used: 98^{o}C for 2 minutes, then 40 cycles of 98^{o}C for 10s, 60^{o}C for 15s and 68^{o}C for 1 minute. For *fuss*, a 2-step PCR was used: 98^{o}C for 2 minutes, then 40 cycles of 98^{o}C for 10s and 68^{o}C for 4 minutes. PCR products were run on a 1% agarose gel.

2.7 Genotyping of *Drosophila* by sequencing

Samples were amplified for sequencing using a single primer (see table in section 2.6.2) and Big Dye Terminator v3.1 Cycle (Thermo Fischer) according to the manufacturer's instructions. Reaction mixtures were cleaned using AgenCourt CleanSEQ magnetic beads according to the manufacturer's instructions, and then sequenced using a ABI 3730 DNA Sequencer (48 Cap). Sequences were analysed using ApE software.

2.8 Measurement of gene expression by qPCR

2.8.1 RNA extraction

Per sample, 10 flies were placed in a FastPrep tube containing Lysis Beads and Matrix D (MP Biomedicals), frozen using dry ice, and stored at -80^{o}C. RNA was isolated using phenol-chloroform extraction based on published methods (Bogart and Andrews, 2006). Briefly, flies were homogenised in 1ml Qiazol (Qiagen) using a Fast Prep 24 machine (6m/s, 40s). Homogenates were incubated at room temperature for 5 minutes and centrifuged (12000rcf, 10 minutes, 4^{o}C). The supernatant was transferred to a fresh RNase-free microcentrifuge tube and 200μl of chloroform (Sigma) added. Solutions were shaken vigorously by hand, incubated at room temperature for 3 minutes, then centrifuged (10000rcf, 15 minutes, 4^{o}C). The upper aqueous phase was transferred to a fresh RNase-free microcentrifuge tube and 0.5ml of isopropanol (Sigma) added. After 10 minutes incubation at room temperature and centrifugation (12000rcf, 10 minutes, 4^{o}C), the supernatant was removed and the pellet washed by inversion with 1ml 75% ethanol. Pellets were left to air-dry for 10 minutes, and then resuspended in 100μl of RNase-free water. Samples were purified using an RNeasy clean-up kit (Qiagen) according to the manufacturer's instructions, including DNase 1 digestion. Final RNA concentrations were measured on a nanodrop.

2.8.2 cDNA preparation

cDNA was created from each RNA sample. Each reaction contained 10μl of 100ng RNA in RNase-free PCR tubes. 2μl of each sample was taken to form the pool from which the -RT control (1μl pool + 9μl water) and +RT control (5μl pool + 5μl water) were made. To each reaction, 1μl of random primer (Promega) at 1/4 dilution, and 1μl of 100mM dNTPs (Bioline) at 1/10 dilution was added. The solutions were heated to 65^{o}C for 5 minutes. After a brief chill on ice, 4μl of 5x first strand buffer (Invitrogen), 2μl of 0.1M DTT (Invitrogen), and 1μl of Superase RNase inhibitor (Ambion) was added to each tube and mixed gently. After 2 minutes incubation at 25^{o}C, 1μl of Superscript II RT was added to all samples except the -RT control (RNase-free water added instead). Reactions were incubated at 25^{o}C for 10 minutes, 42^{o}C for 50 minutes, and 70^{o}C for 15 minutes.

2.8.3 qPCR

For each cDNA sample, 5μl was placed into a well of a 96-well plate. All samples were measured in triplicate and controls in duplicate. 0.2μl of forward primer, 0.2μl of reverse primer, 0.6μl of nuclease-free water, and 6μl of SYBR Green PCR Master Mix (Thermo Fisher) was added to each cDNA well. Plates were sealed with MicroAmp Optical Adhesive Film (Thermo Fisher) and RNA was quantified using an ABI 7900HT Fast Real-Time PCR System (Thermo Fisher). The -RT samples were used as negative controls. RNA amounts were quantified by comparing to serial dilutions of the +RT sample.

2.8.4 Primer design

qPCR primers for genes of interest were designed based on publicly available gene sequences (Attrill et al., 2016). Primers for housekeeping normalisation genes were found in the literature (Ponton et al., 2011). Primers were synthesised by Sigma. All primers used are listed in the following table:

Gene	Primer F Sequence	Primer R Sequence
fuss	tggctcaaccgtcttatcacc	tttcttgcgtaaggctgctg
rab21	atgacataacggaccgagac	tcgctgttcctccaaatcagtc
Mnf	gagcagaagagcccctacct	aatgaaaccctgacgtggac
Rps20	tgtggtgagggttccaagac	gacgatctcagagggcgagt
Tub	tgtcgcgtgtgaaacacttc	agcaggcgtttccaatctg

Chapter 3

Development of a *Drosophila*-based screen to detect energy homeostasis phenotypes

As discussed in chapter 1, there is an unmet need for a high-throughput method to investigate human genetic data related to the regulation of energy homeostasis. *Drosophila* are a genetically tractable model, allowing large numbers of candidate genes to be functionally analysed quickly, economically, and in a spatiotemporal-specific manner.

3.1 Chapter aims

The overall aim of this chapter was to create a high-throughput, *in vivo* screen for energy homeostasis phenotypes using *Drosophila* as a model to study human genes of interest. More specifically the aims were to: identify suitable assays and develop new assays; validate these individual assays using metabolic perturbation; validate the screen as a whole using positive and negative control genes; and explore which of the available *Drosophila* genetic models is the most suitable for use in this screen.

Strategy

3.2 Genetic models available in *Drosophila*

As discussed in chapter 1, one of the advantages of using *Drosophila* as a model organism is the availability of several large collections of lines that can be used for manipulation of the genes of interest. Two different models available from public stock centres were used in this thesis, and are described in more detail in this section: knock-out and knock-down.

3.2.1 The loss-of-function model

Loss-of-function (LoF) flies were obtained from the Bloomington, Kyoto, and Harvard Excelis stock centres. In theory, all of these lines have loss-of-function of the gene of interest (Wangler et al., 2017). Loss-of-function stocks include several types of genetic perturbation, for example: deletion of a large genomic region which includes the gene of interest; mutant alleles induced by chemical mutagens such as EMS; and gene disruptions caused by translocation.

Most of the LoF lines used in this thesis contain a single transposable element (P-, PiggyBac-, or MiET1- element) inserted into or near to the gene of interest to disrupt gene function. Such stocks are available for approximately 70% of *Drosophila* genes (Hummel and Klämbt, 2008).

Male flies were phenotyped. They are smaller and eat less than females (Wong et al., 2009), but males are significantly easier to work with because they do not lay eggs, and so were more suitable for use in a high-throughput screen.

3.2.2 The knock-down model

Two genetic techniques — RNA interference and the UAS_G-GAL4 system — can be used together to produce *in vivo* knock-down of gene expression in precise spatial and temporal patterns. The combination will be referred to in this thesis as RNAi, and it has been shown that this technique generally reduces expression levels of the target gene to 25% of wild-type levels (Perkins et al., 2015; Heigwer et al., 2018).

UAS_G-GAL4 is a binary genetic system adapted from yeast (Brand and Perrimon, 1993). GAL4 protein is a transcriptional activator which recognises and binds to Upstream Activating

Sequences (UAS_Gs) in DNA, resulting in transcription of the gene downstream of the UAS_G sequences. UAS_G-GAL4 does not exist endogenously in *Drosophila* and so can be used to specifically perturb the gene of interest (Duffy, 2002). Transcription of GAL4, and thus expression of the UAS_G-controlled sequence, can be directed by a plethora of different enhancers. Many lines are available to express GAL4 specifically in different tissues, cells, developmental stages, or environmental conditions. The VDRC stock centre alone has 1,000 different GAL4 lines available for purchase. The GAL4 and the UAS_G parts of the system are usually maintained in separate lines, allowing numerous combinatorial possibilities. In this thesis, two GAL4 lines were used: elav-GAL4 which is expressed in the CNS (Robinow and White, 1991) because the brain has a key role in controlling feeding behaviour and energy homeostasis (see section 1.4), and act-GAL4 which is constitutively expressed throughout the body.

RNA interference is a physiological method of targeting mRNAs for degradation. In cells, small double-stranded RNAs with a specific hairpin structure are processed by Dicer to form small interfering (si)RNAs. If the siRNAs are complementary to a gene of interest, they can base pair with the corresponding mRNA and direct specific degradation of this mRNA by the RISC complex. This reduces, but does not eliminate, expression of the gene. By transgenically introducing siRNAs into *Drosophila*, expression of a target gene can be decreased.

The VDRC stock centre has a collection of transgenic fly lines with siRNAs (that are designed to each target a single gene of interest) downstream of inducible UAS_Gs. These lines have been molecularly validated, and shown to provide potent and gene silencing when crossed to GAL4 (Dietzl et al., 2007). There are siRNA lines available to target 91% of *Drosophila* protein-coding genes.

To generate flies for phenotyping, virgin female UAS_G-siRNA flies were crossed to male GAL4 flies and then the male offspring were collected (see figure 3.1).

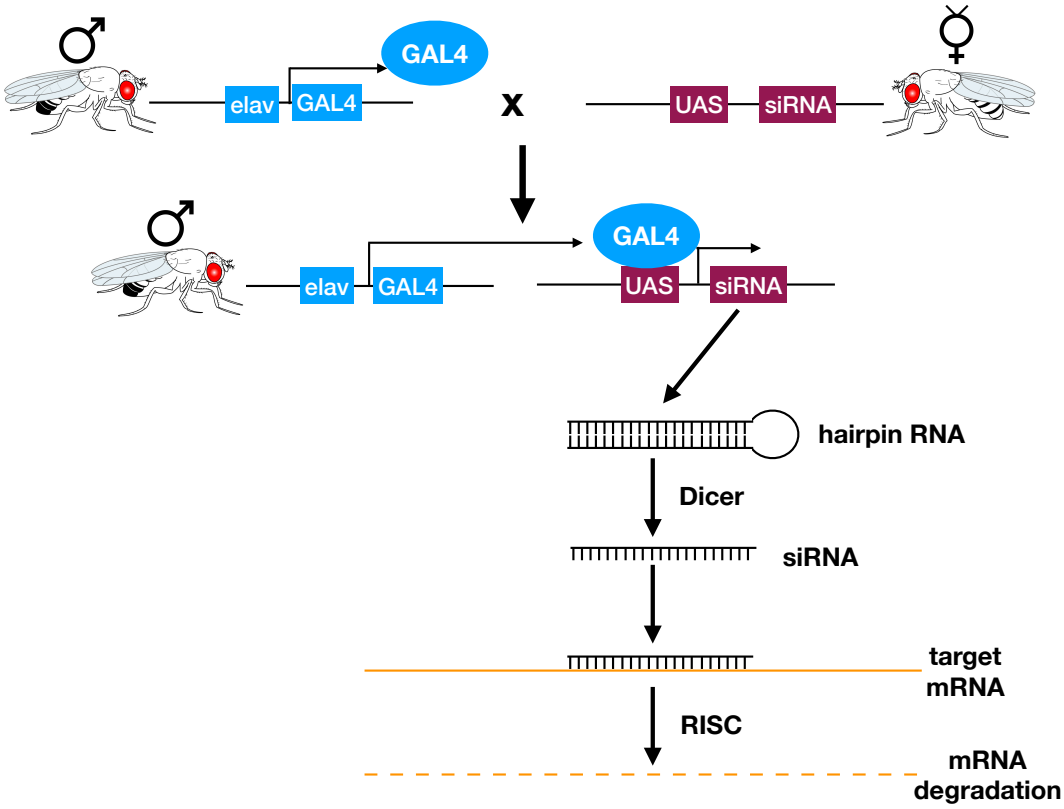


Fig. 3.1 The UAS_G-GAL4 and RNA interference genetic systems can be combined to reduce expression of a specific gene in a spatiotemporal specific manner dependent on the GAL4 promoter.
elav = embryonic lethal abnormal vision, siRNA = small interfering RNA, UAS_G = upstream activating sequence.

3.3 Selection of high-throughput assays to detect energy homeostasis phenotypes

The first step in development of the screen was to identify assays that have been used in published literature to detect energy homeostasis and feeding phenotypes in *Drosophila*, adapted if necessary, and then combined to use as a high-throughput screen. Assays of interest were split into two categories: physiology and behaviour. These assays are described briefly here, and full protocols are detailed in chapter 2.

3.3.1 Assays of physiology

The simplest way to assess flies for existence of an energy homeostasis phenotype is to measure their mass: a chronic excess of energy intake will result in increased storage of energy in the body and thus heavier flies, and vice versa (Jumbo-Lucioni et al., 2010). In this thesis, this assay is referred to as the wet mass assay.

A similar analysis is the dry mass assay: changes in nutrient metabolism can affect the amount of water in a fly, which can be assessed by measuring the desiccated mass of flies compared to the wet mass (Waterson et al., 2014).

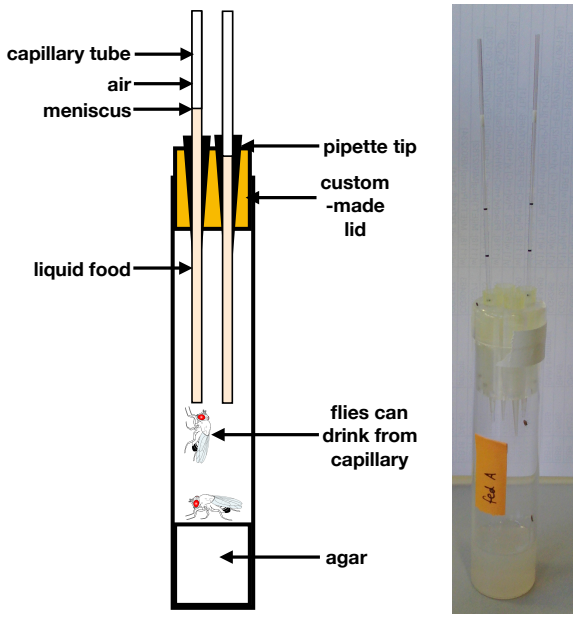
A more direct assessment of the amount of energy stored in a fly is to measure the levels of metabolites present. In particular, the amount of TAG stored in the fat body and the amount of glucose circulating in the haemolymph can both be detected by enzymatic assay of *Drosophila* homogenate (Tennessen et al., 2014; Ugrankar et al., 2015). These two assays are referred to as the TAG assay and the glucose assay, respectively, in this thesis.

Finally, a fly's resistance to starvation is an indirect measure of their levels of energy storage - larger energy stores and the ability to use them efficiently allow a fly to survive for longer when exposed to starvation conditions (Tennessen et al., 2014). This assay is called the starvation resistance assay.

3.3.2 Behavioural assays

The CAFE assay

The Capillary Feeder (CAFE) assay is a method of measuring food intake (Ja et al., 2007). In this assay, the only food available to the flies is liquid food in a capillary tube. The amount of food ingested by flies over a 24 hour period is measured by movement of the food meniscus in the capillary. The CAFE assay provides a measurement of total food intake, which is a combination of both the frequency of feeding and the amount eaten per meal. The set-up of this assay is shown in the following figure.



Dye food absorption assay

Another method of measuring food intake is to mix standard solid *Drosophila* food with food dye (Fast Green FCF dye). Flies are exposed to this food for 24 hours and then homogenised. The amount of dye present in the homogenate (as measured by the absorption of 625nm light) indicates the amount of food eaten (Edgecomb et al., 1994). In this thesis, this assay is called the dye food absorption assay.

Over-feeding dye assay

Using the same dye-containing food, it is also possible to assess acute changes in feeding in response to starvation. In this assay, as shown in figure 3.2, flies are first fasted for 24 hours, and then given normal diet for 20 minutes to feed which is enough time to satiate most wild-type flies (Williams et al., 2014). The flies are then transferred to the dye food which, when ingested, can be easily seen through the cuticle in the abdominal gut using a bench microscope (figure 3.3). Any flies which eat the dye food were not satiated by the 20 minutes of normal food (Edgecomb et al., 1994). This assay — the over-feeding dye assay — measures both satiety and willingness to initiate feeding on a new food.

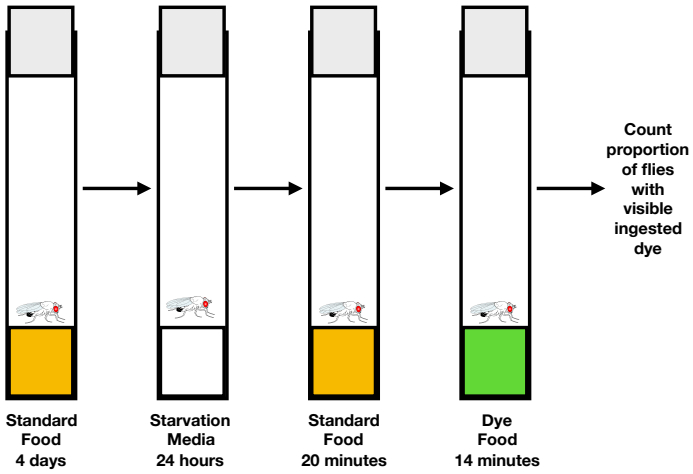


Fig. 3.2 The over-feeding dye assay used to measure food intake by in *Drosophila*

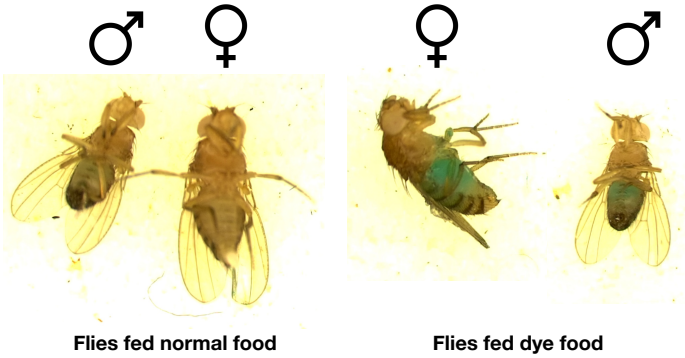


Fig. 3.3 Food dye can be seen through the abdomen after ingestion by *Drosophila*

Negative geotaxis (climbing) assay

In *Drosophila* it has been shown that fasted animals become hyperactive (Yang et al., 2015) whereas flies given high calorie diets decrease their movement (Birse et al., 2010). Therefore movement was included in the screen, as measured using a negative geotaxis climbing assay (Nichols et al., 2012). Flies are put in a column (with no food), tapped to the bottom, and their speed at climbing up the tube is measured (in this thesis, the number above the mid-line within 3s was counted), as shown in figure 3.4.

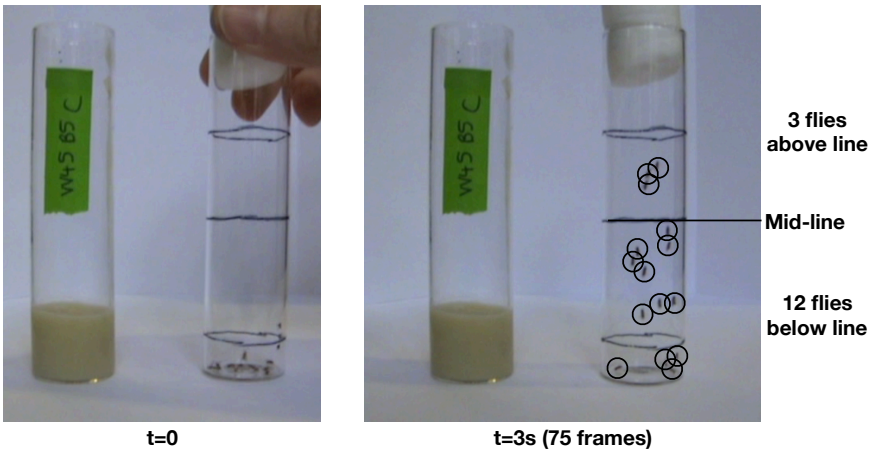


Fig. 3.4 Geotaxic climbing assay used to measure movement of flies

Larval feeding motivation assay

25% of *Drosophila* genes do not produce viable adults when mutant (Wangler et al., 2017) and so cannot be phenotyped using the previously described assays. For many of these genes, homozygous mutants are viable as larvae. Unlike adult flies, young *Drosophila* larvae feed continuously (Zinke et al., 1999). However, upon entering the third instar wandering stage of development, larvae cease constant feeding and so it is at this stage that feeding behaviour is commonly studied (Shen, 2012). The larval assay used here measured feeding motivation (Wu et al., 2003; Melcher and Pankratz, 2005). Briefly, third instar larvae were placed on the edge of an agar plate and the number that moved into the fresh yeast paste in the centre of the plate (hungry larvae) and those that stayed near the edge (non-hungry larvae) was counted after 20 minutes.

Results

3.4 Validation of assays using metabolic perturbation

To validate the protocols for their use as part of a high-throughput screen for energy homeostasis phenotypes, all ten of the physiology and behavioural assays described in section 3.3 were tested for their ability to detect metabolic perturbation in wild-type flies. For the assays using adults, male wild-type (w^{1118}) flies were assayed at five days old having been exposed to either five days of normal diet (ND), five days of high fat diet (HFD), or four days of ND plus one day of fasting. For the larval feeding assay, w^{1118} larvae were either placed in liquid yeast (fed) or water (fasted) for two hours prior to the assay. The results of each assay using these flies/larvae are shown in figure 3.5. Note that the over-feeding dye assay and the starvation assay were not done on fasted flies because the assays themselves involve fasting.

The assays were all able to detect a significant difference in the expected direction between ND and fasted flies/larvae. As would be expected, fasted flies had decreased wet mass, dry mass, TAG and glucose levels, ate more in the CAFE assay and dye absorption assay, and moved faster than ND-fed flies. Fasted larvae also showed increased feeding motivation compared to fed larvae.

Most of the assays were also able to detect a significant difference between ND- and HFD-fed flies. Importantly, for the two assays which could not be tested with fasted flies, the HFD-fed flies show a phenotype in the expected direction: an increase in starvation resistance and a decrease in the over-feeding dye assay. HFD-fed flies also showed a significant increase in TAG levels, decrease in CAFE feeding, and increase in dye food absorption compared to ND-fed flies.

Since all ten of the assays were able to detect at least one of the metabolic perturbations, all were considered suitable for further validation using genetic controls (section 3.7). However, it was expected that fasted and HFD-fed flies would show an opposite phenotype to each other in the assays, but both showed a significant increase in the dye food absorption assay and a decrease in dry mass compared to ND-fed flies, casting some doubt on the accuracy of these two assays.

Starvation resistance data analysis

Data from all of the assays was analysed using Student t-test. Initially the starvation resistance assay data was analysed using a log-rank statistical test as is common in published papers.

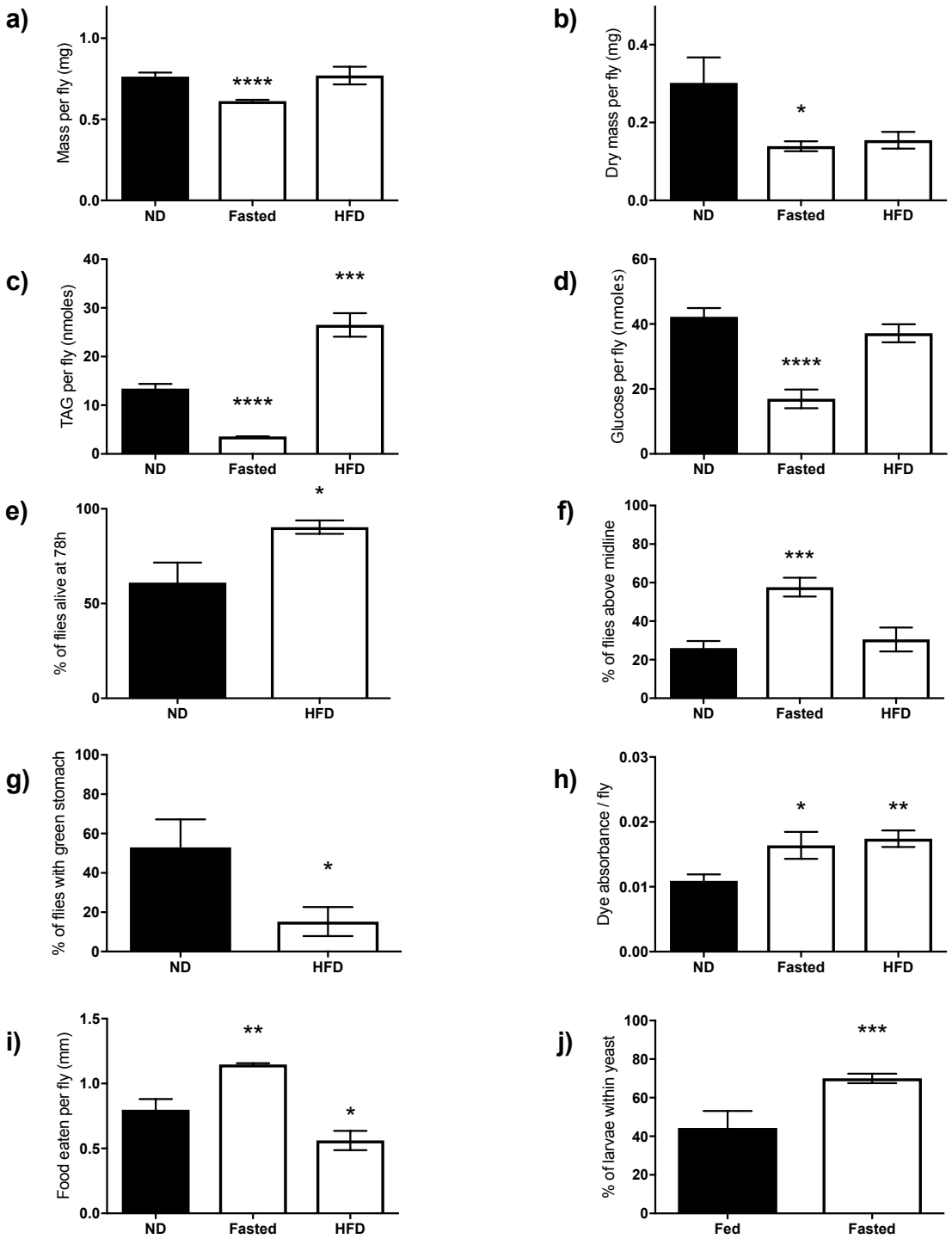


Fig. 3.5 All assays were able to distinguish flies fed on ND from flies subject to 24 hours fasting or 5 days HFD. The physiology assays were a) wet mass, b) dry mass, c) TAG storage, d) glucose storage, and e) resistance to starvation. The behavioural assays were f) geotaxic climbing ability, g) over-feeding dye, h) dye food absorption, i) CAFE feeding j) larval feeding. Mean±SEM is plotted of 7 repeats, each containing (a-h) 15 flies, (i) 8 flies, (j) 20 larvae. Fasted/HFD results were compared to ND results by Student's t-test. *p<0.05, **p<0.01, ***p<0.001, ****p<0.0001.

With this method, the difference in starvation survival between ND- and HFD-fed flies was highly significant ($p<0.0001$). However, this statistical test also assigned 40% (out of 174) of lines in this thesis the same highly significant p value. Thus the test was not useful for discriminating between experimental lines. Instead a Student t-test was used on just one data point: the proportion of flies alive after 78 hours of starvation. This time point was chosen because this is when the survival of w^{1118} flies on ND was closest value to 50% and there was also a significant difference between ND- and HFD-fed w^{1118} flies ($p=0.0222$, figure 3.5e).

The TAG and larval assays

The TAG assay could not be done using RNAi flies because the levels of TAG in the samples were always found to be below the detection limit of the enzymatic assay used. In contrast, most of the LoF lines registered values at least 5 times greater than the detection limit.

The larval feeding assay was difficult to do with the RNAi flies because setting up a cage would require collection of 120 virgin females per line, which is impractical and therefore low-throughput.

All other assays could be used in both the LoF and the RNAi models.

3.5 Development of an assay to detect feeding motivation in adult *Drosophila*

In addition to the high-throughput assays identified in section 3.3, an assay to look for changes in feeding motivation in adult *Drosophila* was developed. As before, w^{1118} flies (5 days old) which had been fasted or fed for 24 hours before the assay were used to develop the protocol.

The set-up of the assay is shown in figure 3.6. Individual flies were put into each lane of a 14-lane chamber, each lane was approximately 0.5cm wide x 0.5cm tall x 10cm long. To move flies into the lanes, they were anaesthetised using ice rather than carbon dioxide (the standard method of immobilising flies) because the latter is a "stress odour" that *Drosophila* avoid and may thus interfere with chemosensory perception of food during the experiment (Siju et al., 2014). Food was placed at one end of each lane. Flies were filmed for one hour and the location of the flies within each lane was detected and tracked using hardware and software developed by Zantiks (www.zantiks.com/). For data analysis, each lane was split into three zones and the total time that each fly spent in each zone during the one hour experiment was calculated.

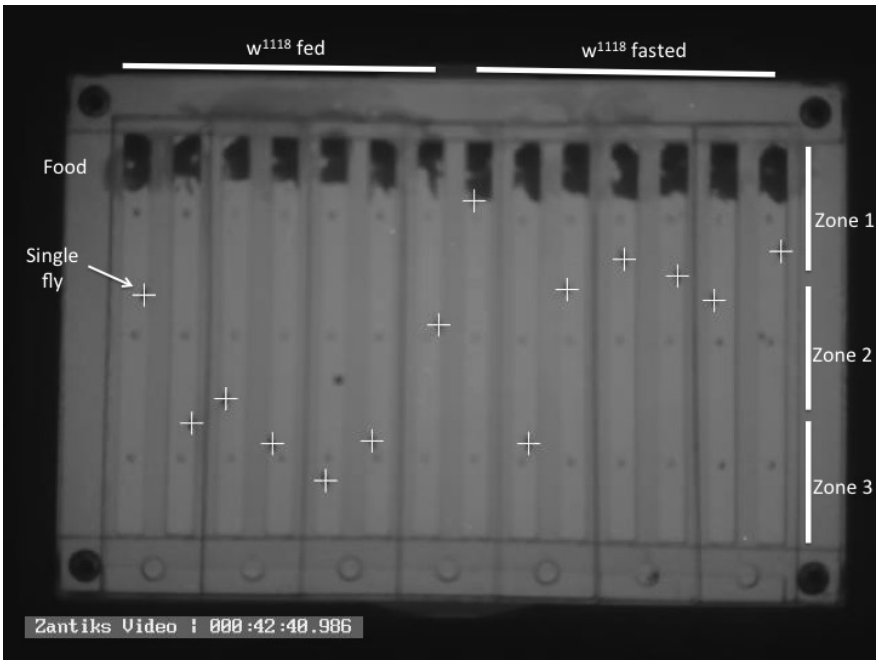


Fig. 3.6 Set-up of the assay for detecting feeding motivation in adult *Drosophila*. The location of each individual fly is marked with a cross.

The first food source placed at the end of each lane was apple juice agar. This food was chosen because its solid nature means that it is easy to place the same amount into each lane. The results are shown in figure 3.7. No significant difference was seen in the behaviour of fasted flies compared to fed flies with this food source.

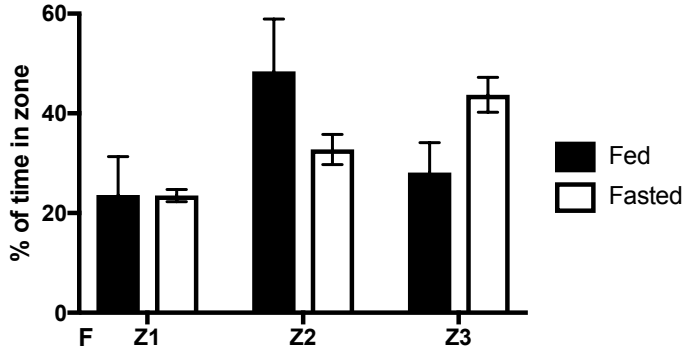


Fig. 3.7 Fasted flies do not spend more time close to the apple juice agar food source than do fed flies. Mean ± SEM is plotted of 7 repeats, each containing 1 fly. Fed and fasted flies were compared by two-way ANOVA. F is the location of the food.

The flies spent approximately equal times in each food zone. Therefore it was hypothesised that the agar did not have a strong enough odour to attract the flies. Thus, the agar was replaced with fresh yeast paste, which is a known attractant for *Drosophila* (Beshel and Zhong, 2013). As a result, the time spent in the food zone did increase, but still no statistical difference was seen between fed and fasted flies (figure 3.8).

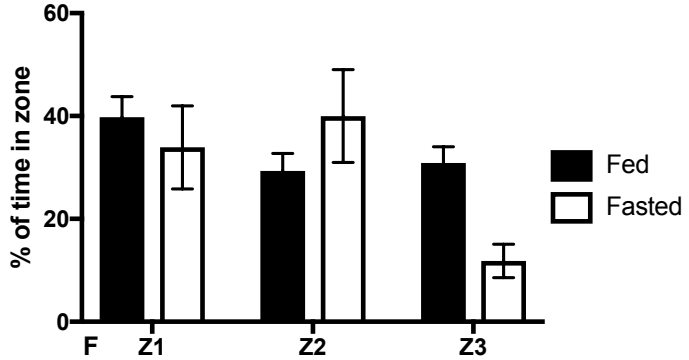


Fig. 3.8 Fasted flies do not spend more time close to the yeast food source than fed flies. Mean ± SEM is plotted of 7 repeats, each containing 1 fly. Fed and fasted flies were compared by two-way ANOVA. F is the location of the food.

One possible explanation for this lack of difference is that the fasted flies feed immediately on the yeast and then behave like fed flies. Alternatively, perhaps the yeast odour is so attractive that the fed flies also approach. To counteract both of these suggestions and ensure that only motivated (hungry) flies approached the food, conditions around the food were made unfavourable using benzaldehyde, which has been shown to be a repellent for *Drosophila* (Vang et al., 2012). The benzaldehyde was placed on cotton wool behind the food.

As can be seen in figure 3.9, this change gave the anticipated results — the fasted flies were more motivated to approach the food than the fed flies, despite the presence of the repellent. However, using the repellent at a dilution of 10^{-4} meant that the flies were strongly repelled and even the fasted flies spent most of their time far from the food source. Therefore, although there was a statistically significant difference, it was not large (p=0.0347, two-way ANOVA, figure 3.9).

Most genetic perturbations are likely to have a smaller effect on feeding motivation than than a severe metabolic perturbation like 24 hours of fasting. Therefore, the sensitivity of the assay was improved by increasing the dilution of benzaldehyde to 10^{-6}. As expected, the fasted flies now spent more time near the food and this greatly increased the significance of the results (p=0.0004, two-way ANOVA, figure 3.9).

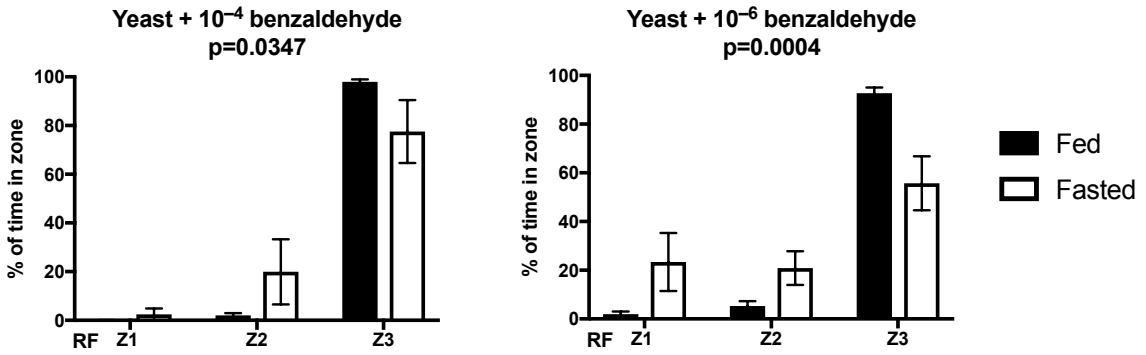


Fig. 3.9 Fasted flies spend more time close to the food source in the presence of repellent than do fed flies. Mean ± SEM is plotted of 7 repeats, each containing 1 fly. Fed and fasted flies were compared by two-way ANOVA. F is the location of the food and R of the repellent.

The requirement for complex equipment and the small number of flies which could be assayed simultaneously made this assay low-throughput. Therefore it was intended for use only with genes already identified as top hits by the high-throughput assays identified in section 3.3.

3.6 Identity of background lines for data analysis

For data analysis, all experimental *Drosophila* lines were statistically compared to lines with closely-matched genetic backgrounds but no perturbation in the genes of interest. As explained in section 3.2, two *Drosophila* models were used — LoF and RNAi — which require different lines to be used as backgrounds.

3.6.1 Background of LoF lines

For the LoF lines used in this thesis, no information about the background was provided by the stock centre for 53% of the lines. The other lines were created from a large number of different wild-type backgrounds including w^{1118}, y^1w^{67c23}, y^1v^1, and ry^{506}, but these lines have been kept separate for many generations, allowing accumulation of modifiers in the genetic background (Busson and Pret, 2007).

Since there is a lot of variety in the background, no single background line is appropriate to use for all of the experimental lines. One solution to this problem is to select a wild-type line to use as background, and then change each LoF line onto this genetic background. To achieve this, lines of interest must be backcrossed to the chosen background strain at least five times to replace the background of the experimental line with that of the control line by recombination. To investigate the practicalities of backcrossing stocks for the screen, and to observe the importance of the background for the phenotypes of interest, three lines (which were obtained for use in chapter 4) were backcrossed to w^{1118}.

The K2 and K3 lines both had a P-element insertion into their target genes (*fur2* and *Ama*) and the background was unknown. The third line (B21) had a known background (y^1v^1), but this was not immediately available to phenotype. B21 contained an siRNA against *twz* inducible by GAL4 (but in the results shown here, the homozygous B21 line was tested i.e. the RNAi was not induced, because the phenotyping was done blinded to genotype).

Backcrossing of these three lines to w^{1118} took approximately four months. All three of the original lines were homozygous viable. During the backcrossing, the K2 line became non-viable as a homozygote. The original K2 line was able to complement this lethality (crossing K2/K2 virgin females to K2*/TM6b males gave viable offspring) suggesting that the *Ama* gene or a nearby gene was mutated during recombination.

For K3 and B21, the effect of changing the background was immediately apparent: the backcrossed flies eclosed three days earlier than the inbred line and produced more than twice as many offspring each generation. In more than half of the assays, at least one of the

inbred lines showed a phenotype (figure 3.10) whereas only one significant phenotype is seen across all the assays in the backcrossed flies.

Therefore it is possible that phenotypes observed in LoF lines could be due to genetic differences in the background of the line, rather than due to perturbation of the gene of interest. However, back-crossing was not high-throughput, and resulted in one of the lines becoming lethal, and so was not deemed an appropriate solution. Further, many of the LoF lines did not have a visible marker (such as eye colour) to track the mutant gene through the backcrossing procedure. Therefore, a different approach to counteract the uncertainty due to background was used: all experimental LoF lines were compared to two backgrounds and the higher (least significant) of the two p values noted. The background lines were w^{1118} and y^1w^1 because they are both widely used in the literature for this purpose.

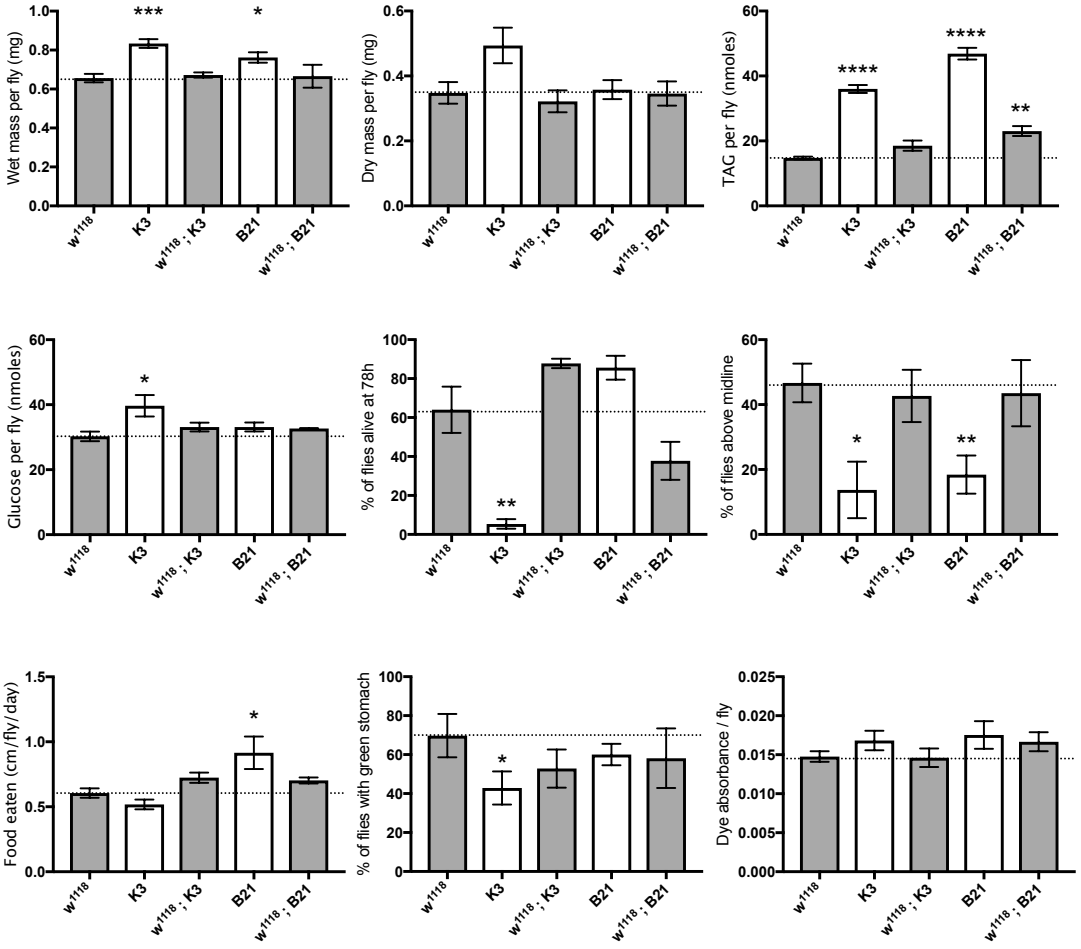


Fig. 3.10 Backcrossed flies (grey) do not have the same phenotypes as inbred flies (white). Mean ± SEM is plotted of 5 repeats. All lines were compared by Student's t-test to w^{1118}. * p<0.05, ** p<0.01, *** p<0.001, **** p<0.0001

3.6.2 Background of RNAi lines

The VDRC stock centre has two UAS_G-siRNA stock collections which were created in different backgrounds: KK and GD. The GD library has P-element based siRNA transgenes inserted at random sites across chromosomes 1, 2, and 3. The KK background contains siRNA transgenes inserted into a chromosome 2 attP site using phiC31 integrase.

The KK and GD backgrounds have slightly different phenotypes to each other and to the elav-GAL4 and act-GAL4/CyO lines (data in appendix A). Since the experimental flies are a cross between the siRNA and GAL4 lines, their background is a mix and so no single one of these parental lines is suitable for use as a background control. Background-matched control flies were created by crossing KK and GD to each GAL4 line to create: KK-elavGAL4, KK-actGAL4, GD-elavGAL4, or GD-actGAL4.

To control for leaky expression of some UAS_G-siRNA transgenes, some papers (for example Beshel et al. (2017)) cross the UAS_G-siRNA flies to wild-type flies. However, this doubles the amount of work required for each line and so is not suitable for a large-scale and high-throughput screen. A preliminary trial phenotyping the homozygous UAS_G-siRNA flies for this purpose was performed (data in appendix A) but did not produce useful information for the heterozygous experimental flies and so was discontinued.

Over-expression of *tiptop* in KK lines

The KK siRNA lines offer three advantages over the GD siRNA lines. Firstly, the siRNA inserts are all at the same site which makes expression more consistent between lines compared to the GD lines where the genomic insertion is random. Secondly, VDRC improved their siRNA design for the KK lines, so increasing the efficiency and specificity of gene disruption compared to the GD lines (Dietzl et al., 2007). Thirdly, the KK lines are easier to work with: the virgins have a more obvious appearance facilitating virgin collection and the larvae do not churn the food to the extent that it drowns many of the virgins.

However, the KK line contained two potential genomic sites for insertion of the UAS_G-siRNA constructs. 25% of the KK lines contain an siRNA insert at the extra site which is located in the 5' untranslated region (UTR) of the gene *tiptop*. These inserts cause ectopic UAS_G-driven expression of *tiptop* and thus non-specific phenotypes with some GAL4 drivers (Green et al., 2014; Vissers et al., 2016).

To test whether our assays were affected by this problem, a line (KK_{tiptop}) with GAL4-responsive UAS_G repeats but no functional siRNA sequence at the *tiptop* site was obtained

(Vissers et al., 2016). This line was lethal when crossed to act-GAL4, likely at the pupal stage because vials contained black pupae. When crossed to elav-GAL4, the offspring flies showed significant increases in the glucose and over-feeding dye assays, and significantly decreased climbing as shown in figure 3.11. They were also heavier and ate more in the CAFE assay, although these results did not reach significance.

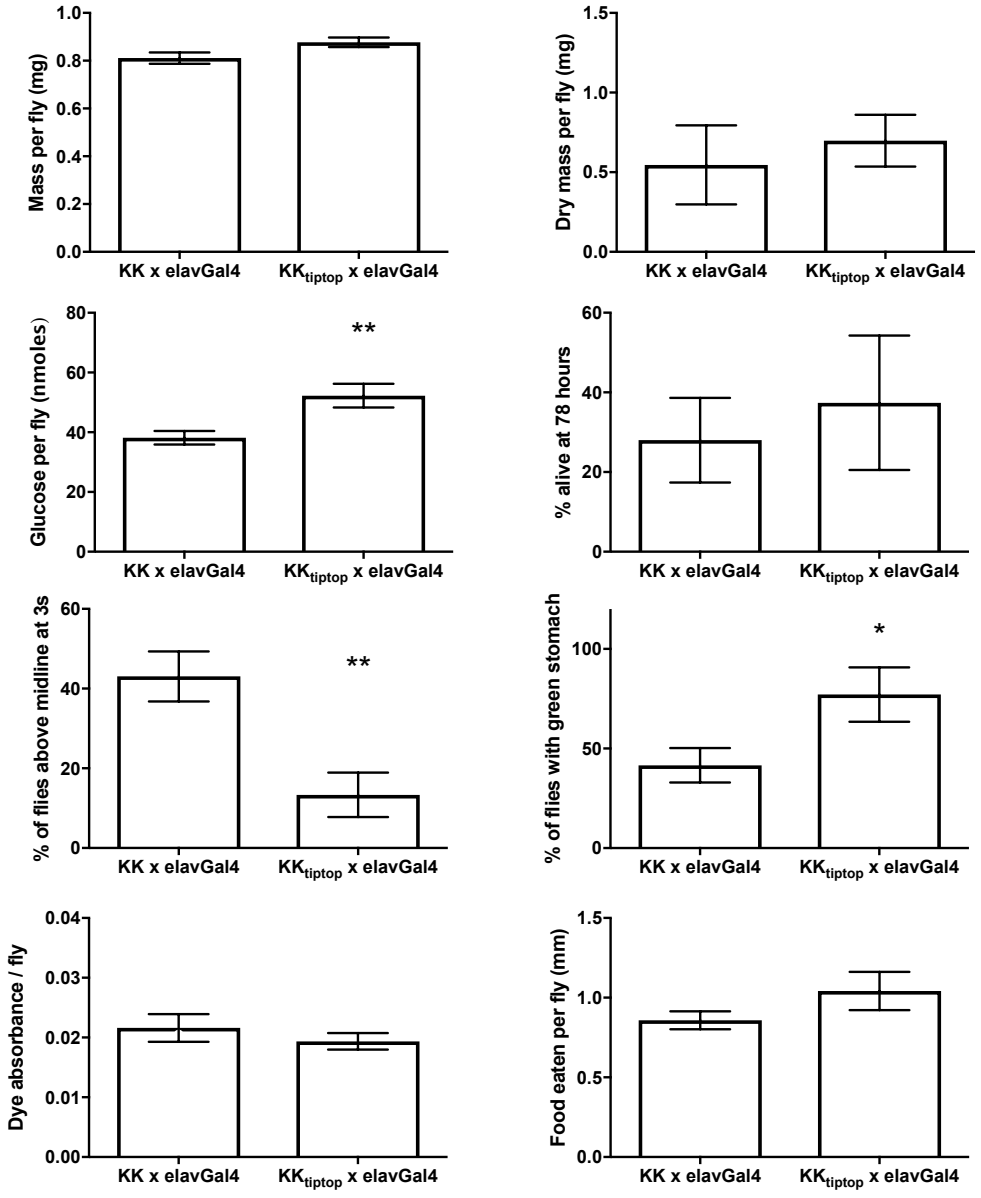


Fig. 3.11 Ectopic UAS$_G$-driven expression of *tiptop* significantly affects results in the glucose, starvation, over-feeding dye, and climbing assays, whilst also increasing mass and CAFE feeding in a non-significant manner. Mean ± SEM is plotted of 5 repeats. Results were compared by Student's t-test. * $p<0.05$, ** $p<0.01$

Since this is a strong overall phenotype, GD lines were always used in preference to KK. If no GD line was available or the GD line(s) had multiple off-target effects, then a KK line was purchased instead. Any KK lines which were lethal before or during the pupal stage when crossed to act-GAL4 were considered as possibly being affected by *tiptop*. The data from such lines were compared to both KK-elavGAL4 and KK_{tiptop}-elavGAL4 as backgrounds.

Compound X GD lines

For the GD lines, if the UAS_G-siRNA insert was on the X chromosome then the stock was delivered from VDRC balanced as a compound X stock, meaning that only male flies carry the UAS_G-siRNA insert. Some published papers have overcome this problem by doing the cross in reverse (using male UAS_G-siRNA and female GAL4 flies) and then phenotyping female offspring (for example, Baranski et al. (2018)). However, since the parental gender of a gene can influence the offsprings' phenotype (Wittkopp et al., 2006), in this thesis the compound Xs were replaced with an FM7 balancer using a series of crosses, allowing the experimental flies to be created in the same manner for all lines. However, this crossing may also have replaced some of the 2nd, 3rd and 4th chromosomes with those of the balancer stock, moving the background of the line away from GD, so any significant phenotypes for these lines should be viewed with caution. Luckily, only 3 lines in this whole thesis were compound X stocks.

3.6.3 Summary of background lines

The different *Drosophila* models used in this thesis (LoF, whole body RNAi and neuronal RNAi) each require different lines to be used as background for statistical comparison. These are summarised in the following table:

Model	**Background**	**Background line used**
LoF	variable	yw and w^{1118}
Whole body RNAi	GD	GD-actGAL4
	KK	KK-actGAL4
Neuronal RNAi	GD	GD-elavGAL4
	KK (viable whole body)	KK-elavGAL4
	KK (lethal whole body)	KK-elavGAL4 and KK_{tiptop}-elavGAL4

where GD is w^{1118}, and KK is $y^{-}w^{1118}$;P{attP,$y^{+}w^{3}$}VIE-260B.

Development and metabolism are influenced by environmental conditions such as humidity, temperature, and food (Strassburger and Teleman, 2016) which could potentially lead to false positive results. Therefore, all experimental flies were kept in atmosphere-controlled incubators, and every line was phenotyped over several different batches. Development and metabolism are also affected by environmental crowding and so flies were raised at approximately constant density. In addition, experimental flies were allowed to mate before being phenotyped because mating alters gene expression, metabolic parameters and food intake (Ellis and Carney, 2010; Kubli, 2010; Papanastasiou et al., 2013). Finally, the assays were always performed on the same day relative to eclosion, and at a certain time of day, because *Drosophila* feeding behaviour and activity levels are affected by age and circadian rhythms (Xu et al., 2008; Iliadi and Boulianne, 2010).

3.7 Validation of the high-throughput screen using positive and negative control genes

As shown in section 3.4, all of the high-throughput physiological and behavioural assays identified in section 3.3 were able to detect differences in flies with diet-induced metabolic perturbations. To further validate the screen for large-scale use to detect energy homeostasis phenotypes, a series of positive and negative control genes was tested. Phenotyping of these control genes served several purposes.

1. It allowed each assay to be tested as to whether it is high-throughput enough to use in a large-scale screen — all assays were found to meet this criterion.

2. It allowed the precision and accuracy of each assay to be assessed (see sections 3.7.6 and 3.7.7) which in turn enabled a system to be developed to aid data inspection by assigning a single score value to each gene (see section 3.8).

3. Comparison of the negative control genes to the positive control genes allowed the threshold for a result to be considered positive to be established (see section 3.8).

4. The LoF and RNAi *Drosophila* models could be compared and contrasted for their suitability for use in the screen (see section 3.10.6).

3.7.1 Identification of positive control genes

Nine *Drosophila* genes were selected from published literature to use as positive controls for the screen (table 3.1). These genes have all been shown to have a phenotype in at least one of the screen assays, and many are involved in the neuronal control of feeding and were discussed in section 1.11. In humans, some of these genes are associated with Mendelian metabolic disorders, for example leptin and monogenic obesity (see chapter 1). Others have been implicated in polygenic diseases by GWAS, for example TFAP2B and obesity (Speliotes et al., 2010).

In addition, a mammalian gene from internal lab group data was included: PTPN2. A literature search returned no published studies of energy homeostasis phenotypes for the fly orthologue of this gene, *Ptp61f*. The purpose of including this gene was proof-of-principle (i.e. is the screen able to detect this gene?) since the intention of the screen was to use *Drosophila* to identify novel genes important in mammals. PTPN2 was first identified by GWAS in dogs, and sequencing of this locus found a 4bp insertion in the 5'UTR of PTPN2 that is significantly associated with adiposity (Eleanor Raffan, personal communication).

LoF and/or RNAi lines of these positive control genes were ordered from publicly available stock centres.

Table 3.1 Genes used as positive controls

Drosophila gene	Human gene	Reference
upd1	leptin	Beshel et al. (2017)
upd2	leptin	Rajan and Perrimon (2012)
DSK	CCK	Söderberg et al. (2012)
bmm	ATGL	Gronke et al. (2005)
chico	IRS1	Bohni et al. (1999)
leucokinin	tachykinin	Al-Anzi (2010)
hugin	neuromedin U	Melcher and Pankratz (2005)
twz	KCTD15	Williams et al. (2014)
TfAP-2	TFAP2B	Williams et al. (2014)
Ptp61F	PTPN2	Eleanor Raffan, personal communication

3.7.2 Negative control genes

Identification of neuronal disease negative control genes

For negative controls, 11 genes were chosen that have been associated with human diseases that have a neuronal etymology (because of the key role that the CNS has in controlling energy homeostasis) but whose primary symptoms are non-metabolic. Some of the genes have Mendelian links with their respective disease, and the others have been associated by GWAS. In this thesis, these 11 genes are referred to as the neuronal disease negative control genes, and they are listed in table 3.2. However, a literature search showed that all of these neuronal diseases have been associated with weight loss or gain in patients.

Identification of peripheral disease negative control genes

Therefore, another set of negative control genes were selected from diseases which do not have a strong neuronal effect — the peripheral disease negative control genes (see table 3.3). No published data was found showing that these diseases lead to weight gain/loss, but it should be noted that the risk of many of these diseases is increased by being over- or under-weight.

3.7.3 Orthologues of control genes

Drosophila orthologues of the negative control genes and of PTPN2 were identified using the ENSEMBL orthologue tool (Zerbino et al., 2018). All genes had sequence homology $>20\%$, and most were $>40\%$.

According to published microarray data (Graveley et al., 2011; Gramates et al., 2017), all of these *Drosophila* genes are expressed in the larval and/or adult CNS.

Genes were randomly assigned to be assayed by either RNAi or LoF, or both. Phenotyping was done in a manner blinded to genotype. It was expected that the neuronal disease negative control genes would have a stronger phenotype than the peripheral disease negative controls, but these would both have a distinguishably weaker phenotype than the positive control genes.

Table 3.2 **Neuronal disease negative control genes**

Disease	Patients experience	Gene reference	Human gene	*Drosophila* gene	Homology (%)
Alzheimer's	weight loss (White et al., 1996)	Kelleher et al. (2017)	PSEN1	*psn*	51
		Xu et al. (2015)	PICALM	*lap*	41
Parkinson's	weight loss (Kashihara, 2006)	Lill (2016)	MCCC1	*CG2118*	52
			GAK	*aux*	39
			MAPT	*tau*	20
Autism	overweight (Hill et al., 2015)	Jonsson et al. (2014)	SHANK3	*prosap*	26
Depression	weight loss/gain (Weissenburger et al., 1986)	Huo et al. (2016)	SLC25A37	*mfrn*	46
Multiple cognitive	multiple	Sinkus et al. (2015)	CHRNA7	*nAChRα5*	46
Schizophrenia	obesity (Panariello et al., 2011)	Pantavou et al. (2016)	FZD3	*fz*	42
Huntington's	weight loss (Carroll et al., 2015)	Young et al. (2012)	HIP14	*hip14*	44
ALS	weight loss (Holm et al., 2013)	van Rheenen et al. (2016)	UNC13A	*unc13 / unc13-4a*	56

Table 3.3 **Peripheral disease negative control genes**

Disease	At increased risk	Gene reference	Human gene	*Drosophila* gene	Homology (%)
RLS	high adiposity (Gao et al., 2009)	Winkelmann et al. (2011)	MEIS1	*hth*	60
			BTBD9	*btbd9*	50
Atopic dermatitis	overweight (Zhang and Silverberg, 2015)	Paternoster et al. (2015)	MRPS21	*mRpS21*	49
Allergic Disease	high BMI (Lokaj-Berisha et al., 2015)	Ferreira et al. (2017)	GLB1	*ect3*	43
HIV infection	-	Johnson et al. (2015)	KDM4C	*kdm4a*	48
			UBR4	*poe*	35
Osteoporosis	anorexic patients (Grinspoon et al., 2000)	Zheng et al. (2015)	EN1	*en*	36

3.7.4 Genetic control lines affected by over-expression of *tiptop*

As discussed in section 3.6.2, some KK lines are affected by over-expression of *tiptop* and this causes a strong phenotype in the assays used here. Of the genetic control lines assayed by RNAi, eight had a KK background and four of these were lethal when crossed to actGAL4. These four lines were therefore compared to both KK-elavGAL4 and KK_{tiptop}-elavGAL4 backgrounds, as shown in figure 3.12.

	Line	Background	Fly Gene	Human Gene	Mass	Glucose	Starvation	CAFE feeding	Dye feeding	Climbing
Peripheral negative controls	VV32	KK	*hth*	MEIS1	0.5951	**0.0425**	0.1009	0.7935	0.0534	0.2174
		tiptop			0.3059	0.8621	0.0667	0.1237	0.672	0.1024
	VV75	KK	*mRpS21*	MRPS21	0.3154	**0.0095**	**0.0124**	0.9258	0.8996	0.0623
		tiptop			**0.0113**	0.3962	0.0877	**0.0192**	**0.0324**	0.4468
Neuronal negative control	VV1	KK	*psn*	PSEN1	0.5698	**0.0061**	0.2605	0.105	**0.0473**	0.2174
		tiptop			0.1288	0.702	0.5968	0.5066	0.7914	0.1024
Positive control	VV54	KK	*twz*	KCTD15	0.4272	0.2527	0.4445	0.1705	0.3336	0.9324
		tiptop			**0.0077**	**0.0047**	0.8444	0.1783	**0.0202**	**0.0385**

Fig. 3.12 p values of potential *tiptop*-affected lines compared to both KK-elavGAL4 and KK_{tiptop}-elavGAL4 backgrounds. The background with the smaller phenotype was chosen, indicated in orange. Red is an increase and blue is a decrease compared to background, and significant results ($p<0.05$) are highlighted

For the *hth* and *psn* siRNA lines, there were significant phenotypes when compared to KK-elavGAL4, but not when compared to KK_{tiptop}-elavGAL4 , and thus these lines were assigned to a KK_{tiptop} background.

Conversely, *twz* showed a strong phenotype when compared to KK_{tiptop}-elavGAL4, but not when compared to KK-elavGAL4 and so this line was assigned to a KK background. Although *twz* is a positive control, choosing the stronger phenotype would be wrong because (i) the data should be treated consistently across the positive and negative control genes to avoid bias in the results, and (ii) the strong phenotype is the opposite direction to that reported in the literature (Williams et al., 2014).

Finally, for *mRpS21* the choice of background was less clear as both gave similar numbers of significant phenotypes. KK was chosen because the average p value across all six assays was greater (less significant).

3.7.5 Phenotypes of control genes

The results of the screen for the positive and negative control genes is shown in figures 3.13 to 3.21. In all figures, results from LoF lines were normalised to w^{1118}, and statistically compared to both w^{1118} and yw. Results from whole body RNAi were normalised and statistically compared to either KK-actGAL4 or GD-actGAL4 as appropriate. Results from neuronal RNAi were normalised and statistically compared to either KK-elavGAL4, GD-elavGAL4, or KK_{tiptop}-elavGAL4 as appropriate.

Wet mass

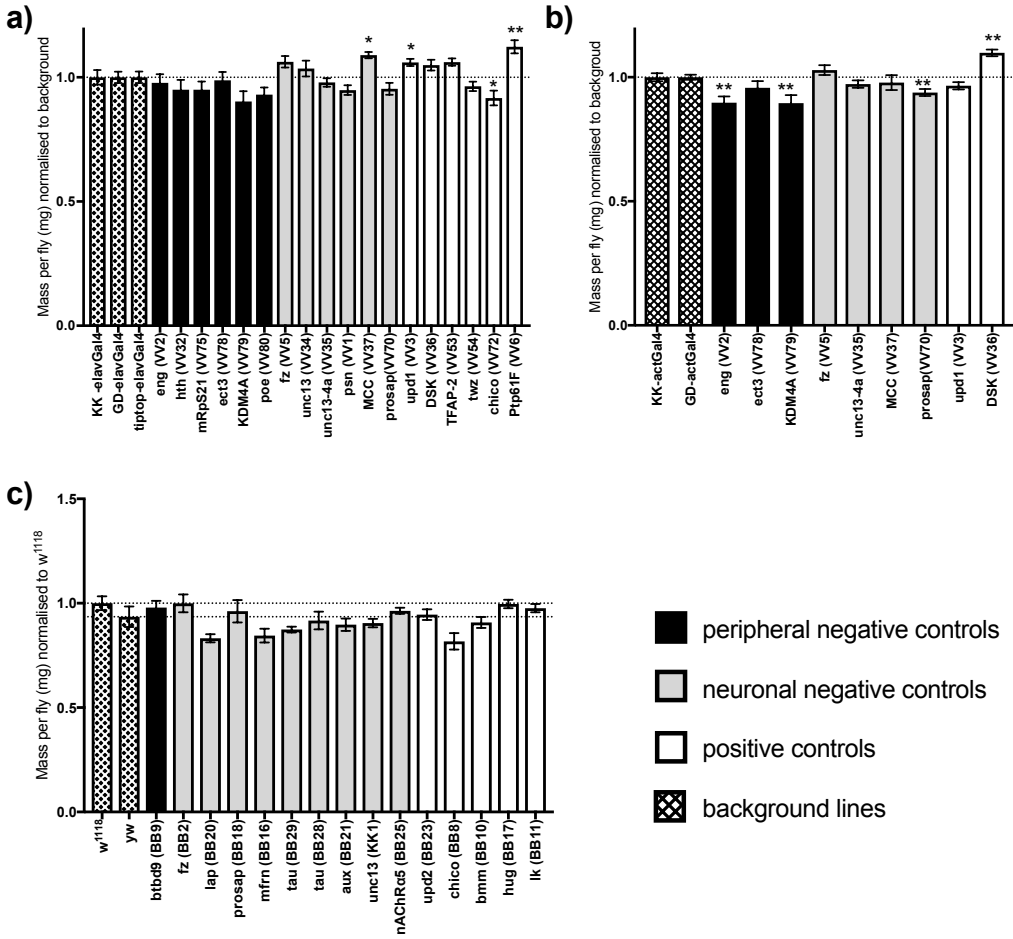


Fig. 3.13 Wet mass of flies with control genes disrupted. Genes were perturbed by (a) neuronal RNAi, (b) RNAi in the whole body, and (c) LoF. Mean ± SEM is plotted of (a, b) 5 repeats or (c) 7 repeats, each containing 15 flies. Results were compared to appropriate background line by Student's t-test. * $p<0.05$, ** $p<0.01$

Dry mass

Not all control genes were phenotyped with the dry mass assay, for reasons explained in section 3.7.7.

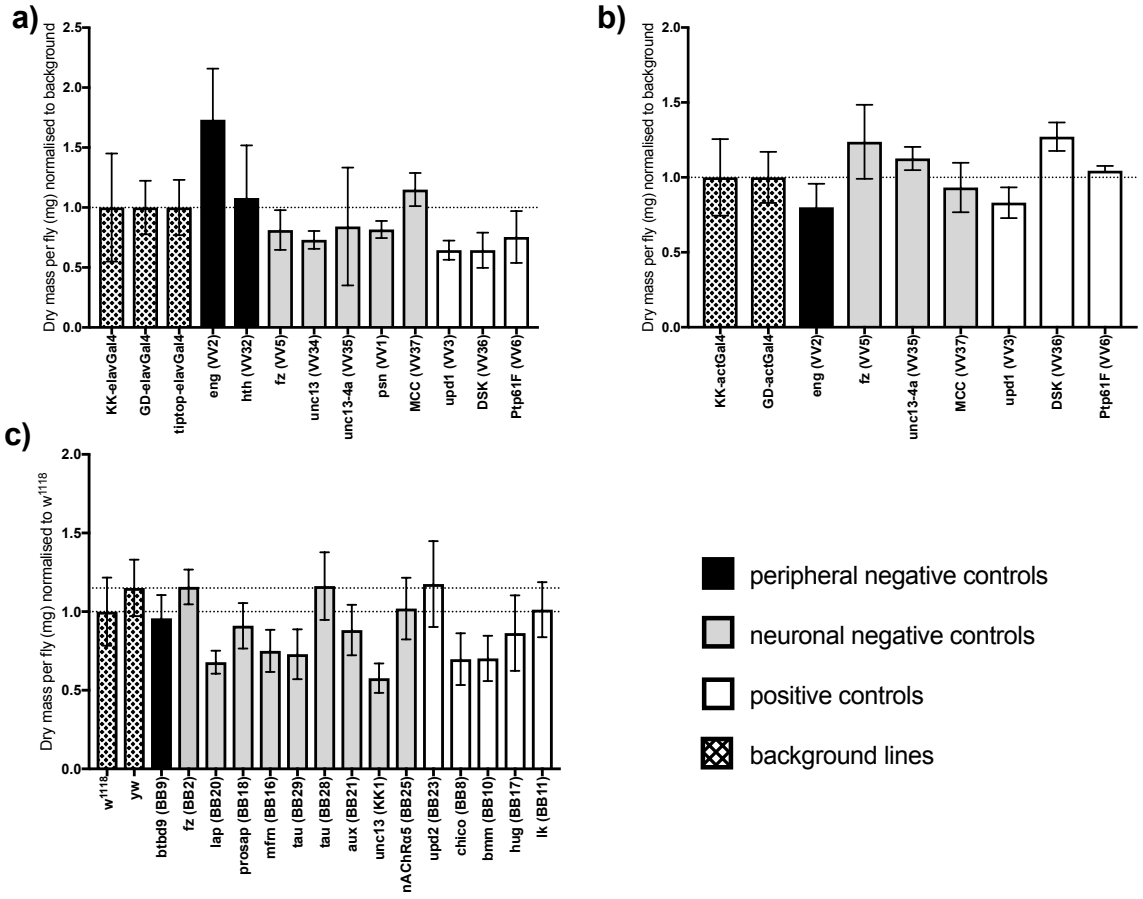


Fig. 3.14 Dry mass of flies with control genes disrupted. Genes were perturbed by (a) neuronal RNAi, (b) RNAi in the whole body, and (c) LoF. Mean ± SEM is plotted of (a, b) 5 repeats or (c) 7 repeats, each containing 15 flies. Results were compared to appropriate background line by Student's t-test.

TAG

As explained in section 3.4, only the LoF lines were phenotyped with the TAG assay.

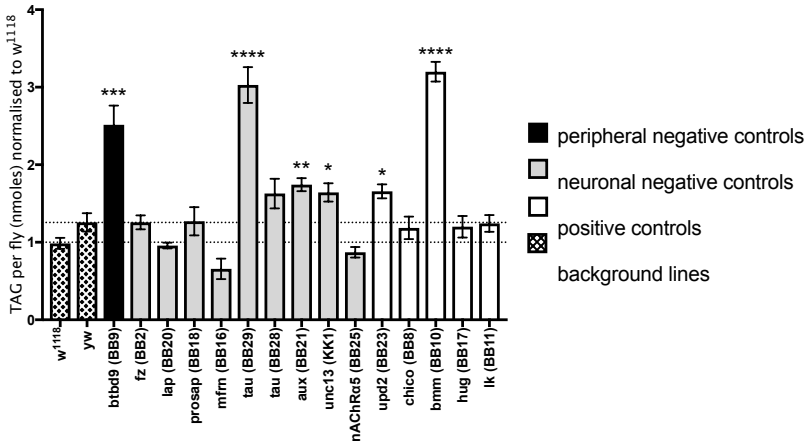


Fig. 3.15 TAG levels in lines with LoF of control genes. Mean ± SEM is plotted of 7 repeats, each containing 15 flies. Results were compared to appropriate background line by Student's t-test. * p<0.05, ** p<0.01, *** p<0.001, **** p<0.0001

Glucose

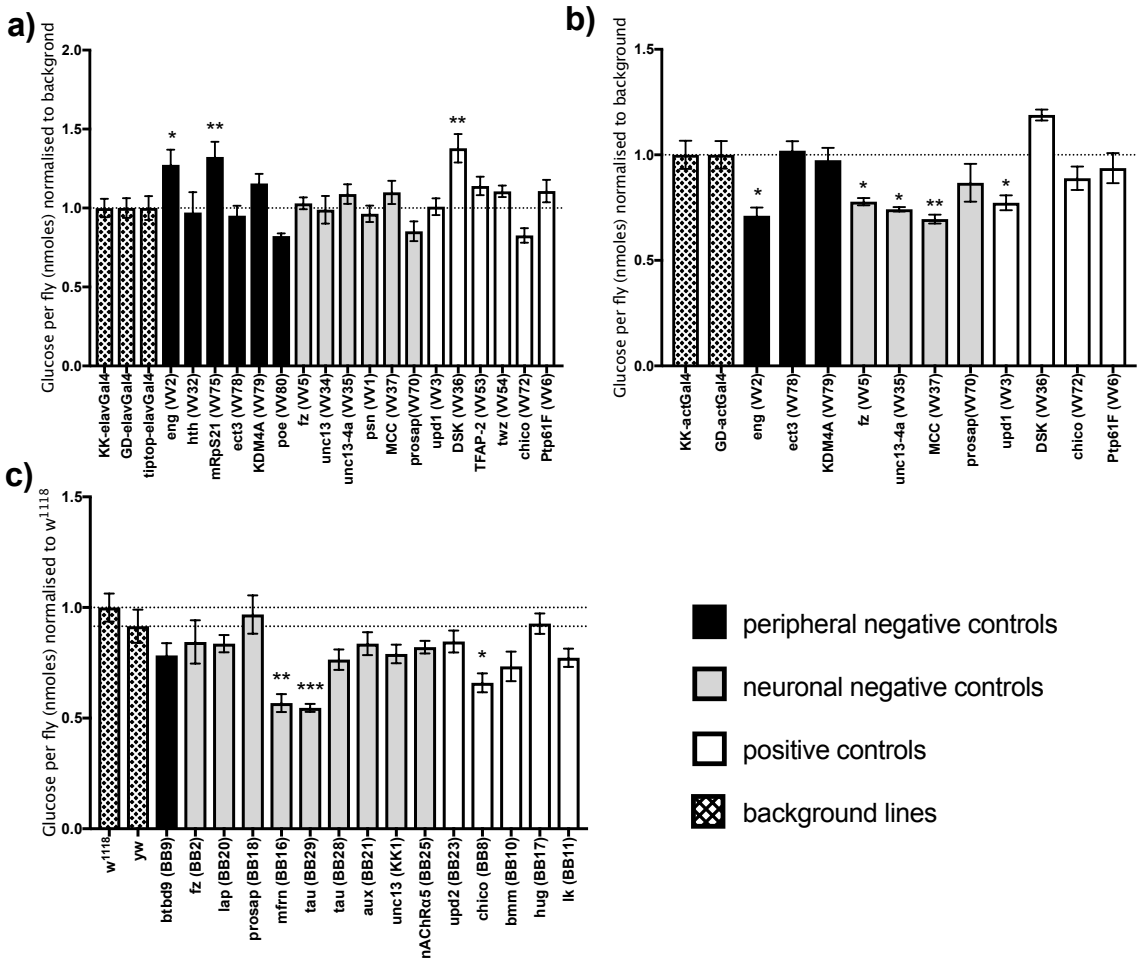


Fig. 3.16 Glucose levels in flies with control genes disrupted by (a) neuronal RNAi, (b) RNAi in the whole body, and (c) LoF. Mean ± SEM is plotted of (a, b) 5 repeats or (c) 7 repeats, each containing 15 flies. Results were compared to appropriate background line by Student's t-test. * $p<0.05$, ** $p<0.01$, *** $p<0.001$

Starvation resistance

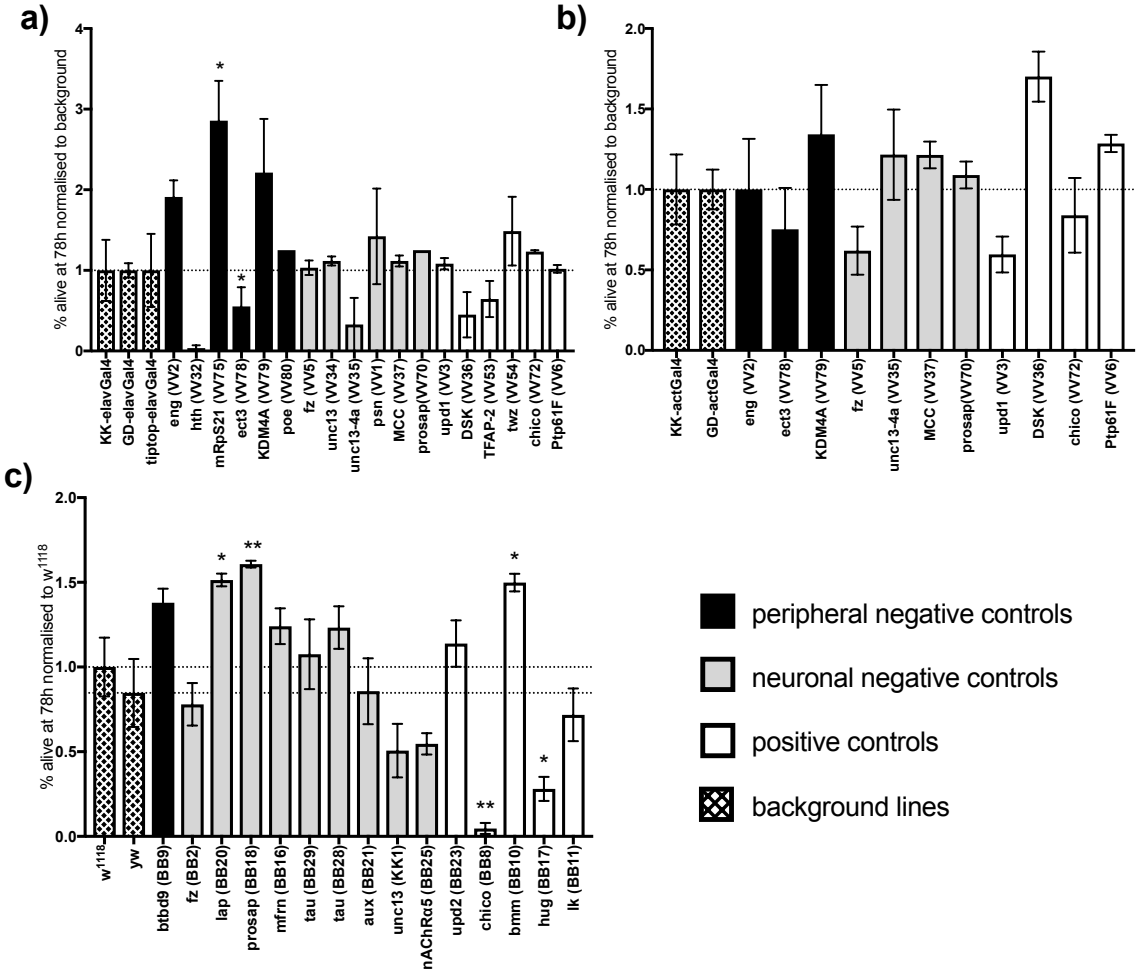


Fig. 3.17 Resistance to starvation of flies with control genes disrupted. Genes were perturbed by (a) neuronal RNAi, (b) RNAi in the whole body, and (c) LoF. Mean ± SEM is plotted of (a, b) 5 repeats or (c) 7 repeats, each containing 15 flies. Results were compared to appropriate background line by Student's t-test. * $p<0.05$, ** $p<0.01$

CAFE assay

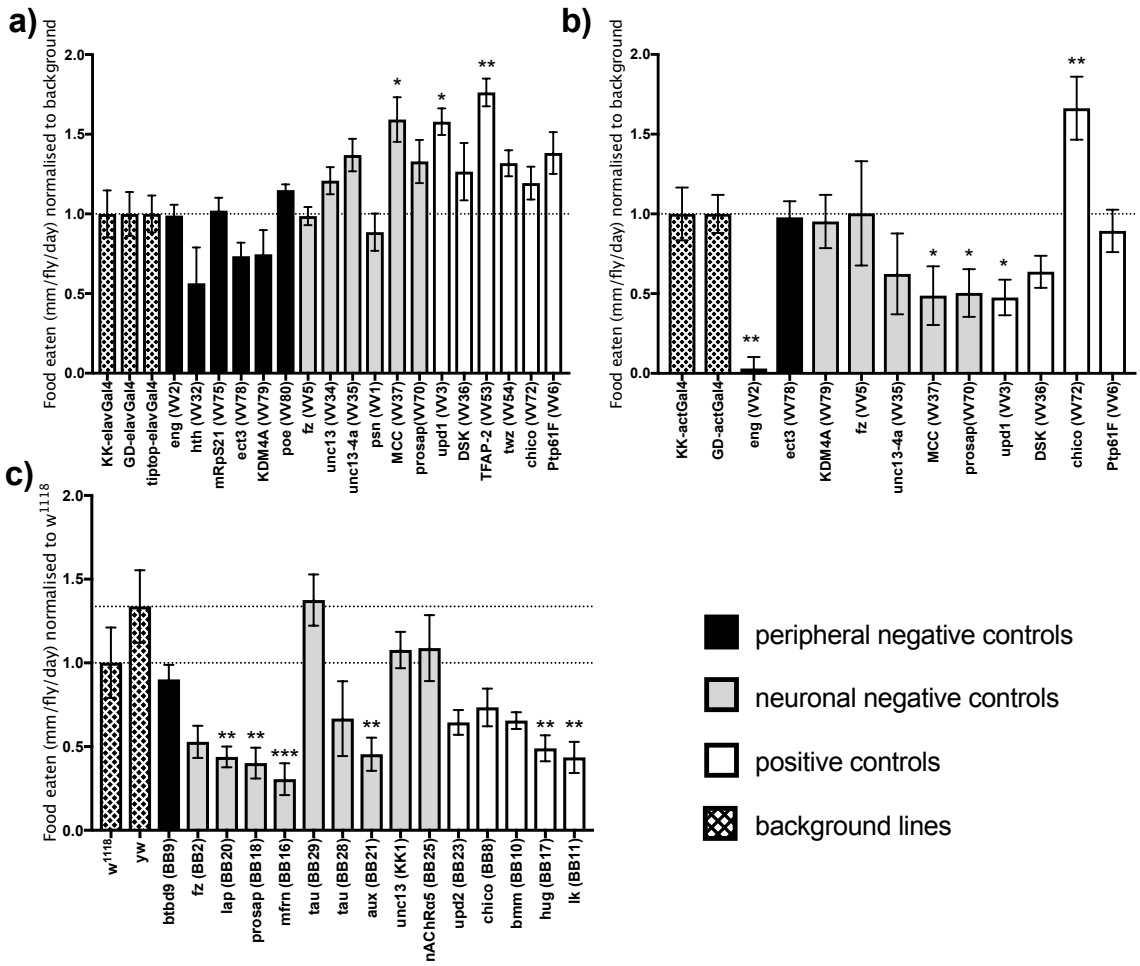


Fig. 3.18 Food intake, as measured by the CAFE assay, by flies with control genes disrupted by (a) neuronal RNAi, (b) RNAi in the whole body, and (c) LoF. Mean ± SEM is plotted of (a, b) 5 repeats or (c) 7 repeats, each containing 8 flies. Results were compared to appropriate background line by Student's t-test. * $p<0.05$, ** $p<0.01$, *** $p<0.001$

Over-feeding dye assay

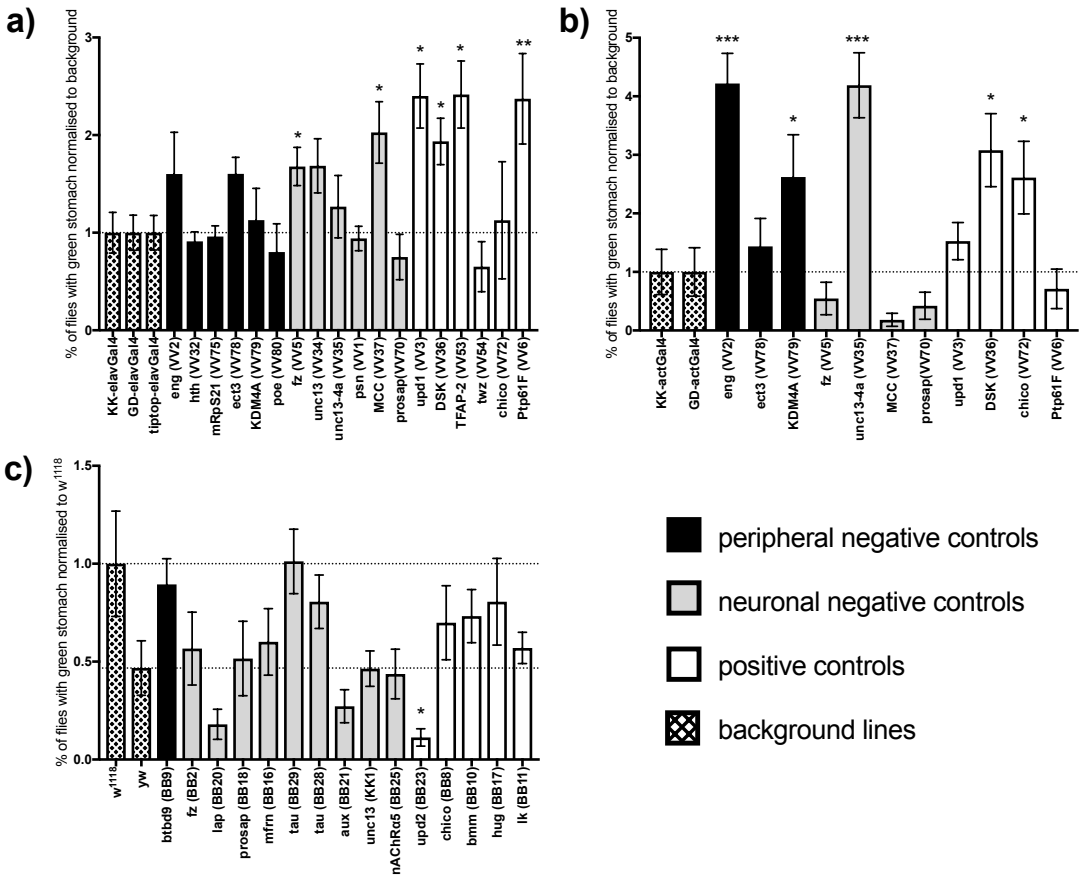


Fig. 3.19 Ingestion of food after starvation, as assessed by the appearance of dye in the abdomen of flies, by flies with control genes disrupted. Genes were perturbed by (a) neuronal RNAi, (b) RNAi in the whole body, and (c) LoF. Mean $\pm$ SEM is plotted of (a, b) 5 repeats or (c) 7 repeats, each containing 15 flies. Results were compared to appropriate background line by Student's t-test. * $p<0.05$, ** $p<0.01$, *** $p<0.001$

Dye food absorption assay

Not all control genes were phenotyped with the dye food absorption assay assay, for reasons explained in section 3.7.7.

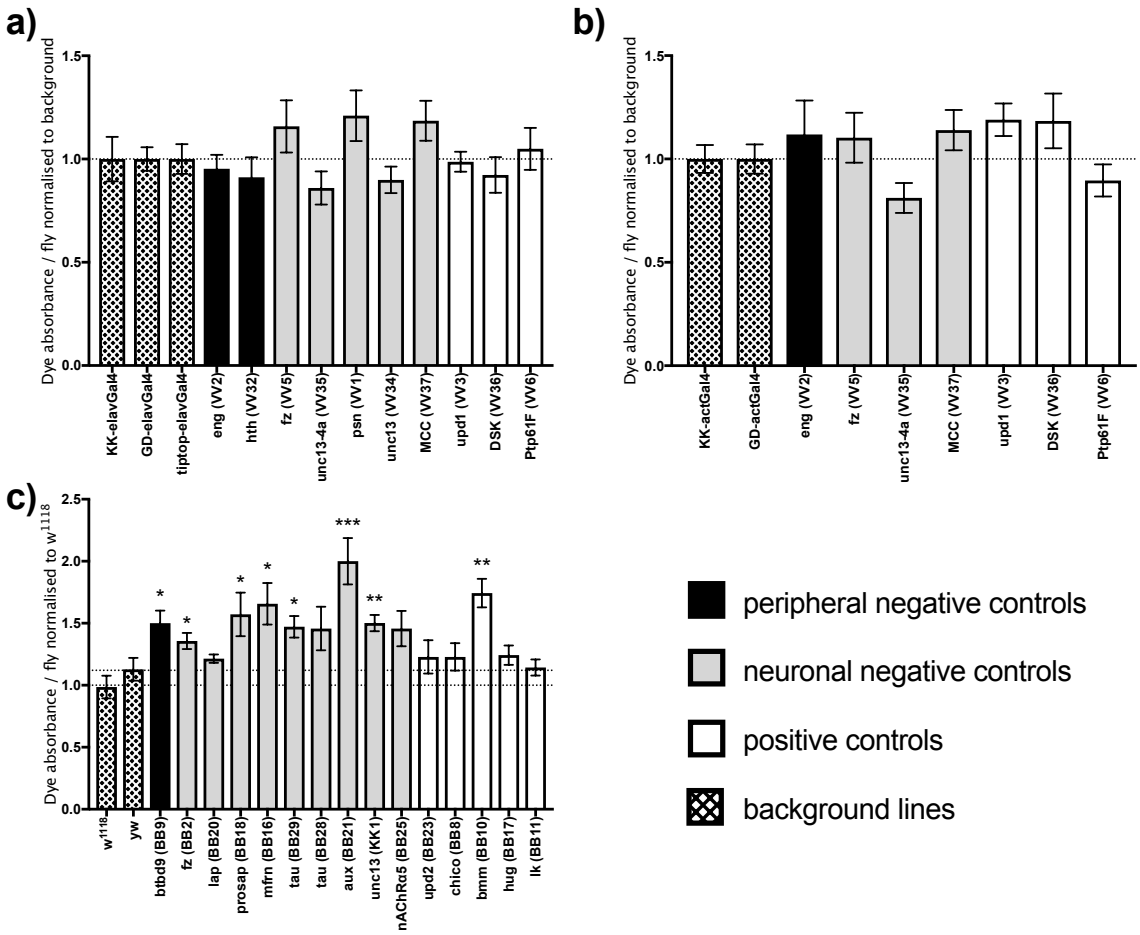


Fig. 3.20 Ingestion of food, as measured by the absorbance of food dye, by flies with control genes disrupted. Genes were perturbed by (a) neuronal RNAi, (b) RNAi in the whole body, and (c) LoF. Mean ± SEM is plotted of (a, b) 5 repeats or (c) 7 repeats, each containing 8 flies. Results were compared to appropriate background line by Student's t-test. * $p<0.05$, ** $p<0.01$, *** $p<0.001$

Negative geotaxis climbing assay

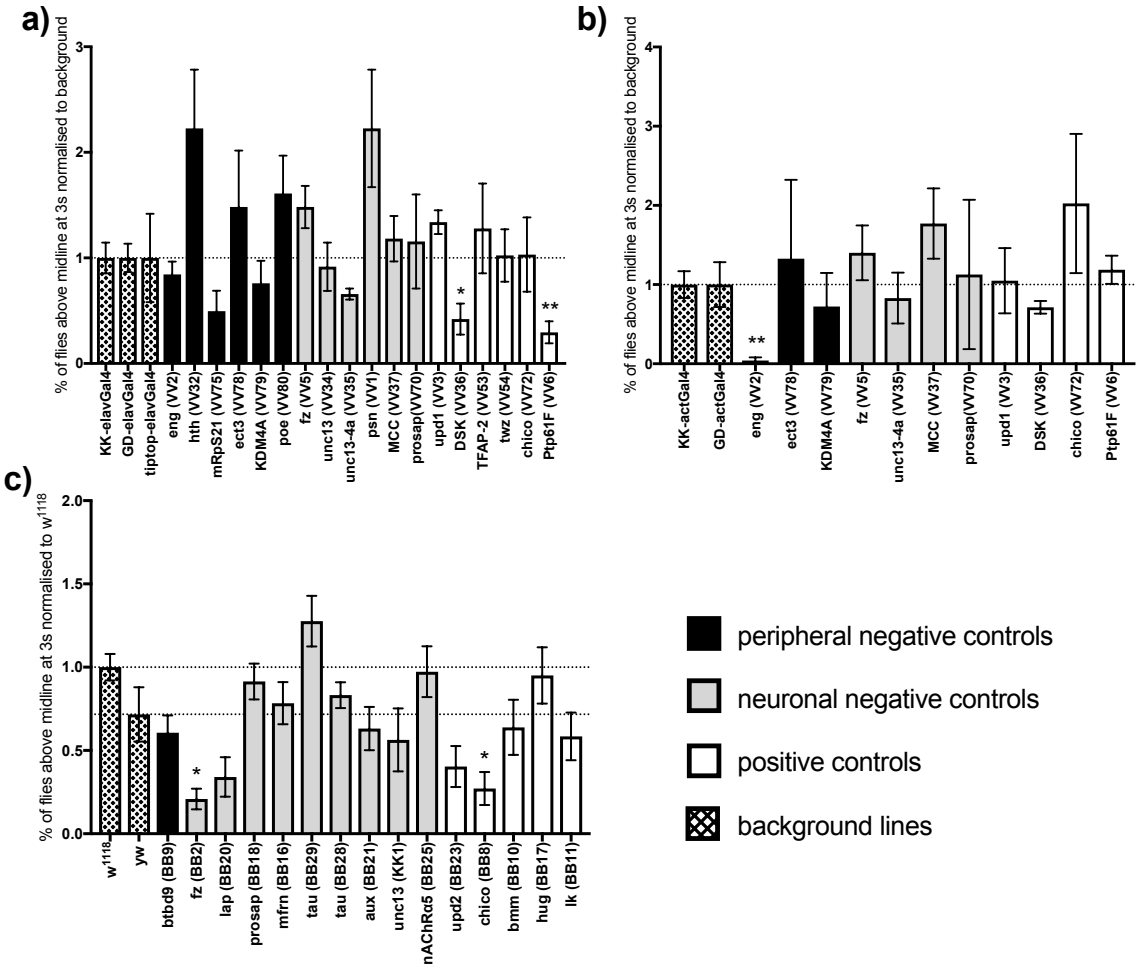


Fig. 3.21 Climbing ability of flies with control genes disrupted. Genes were perturbed by (a) neuronal RNAi, (b) RNAi in the whole body, and (c) LoF. Mean $\pm$ SEM is plotted of (a, b) 5 repeats or (c) 7 repeats, each containing 15 flies. Results were compared to appropriate background line by Student's t-test. * $p<0.05$, ** $p<0.01$

Larval feeding assay

The larval feeding assay was only performed on the LoF lines, as explained in section 3.4. Figure 3.22 shows the results of the assay done on larvae in a fed or fasted state. Figure 3.23 shows the change in behaviour between the two states - genes which show no difference are responding abnormally to fasting/feeding.

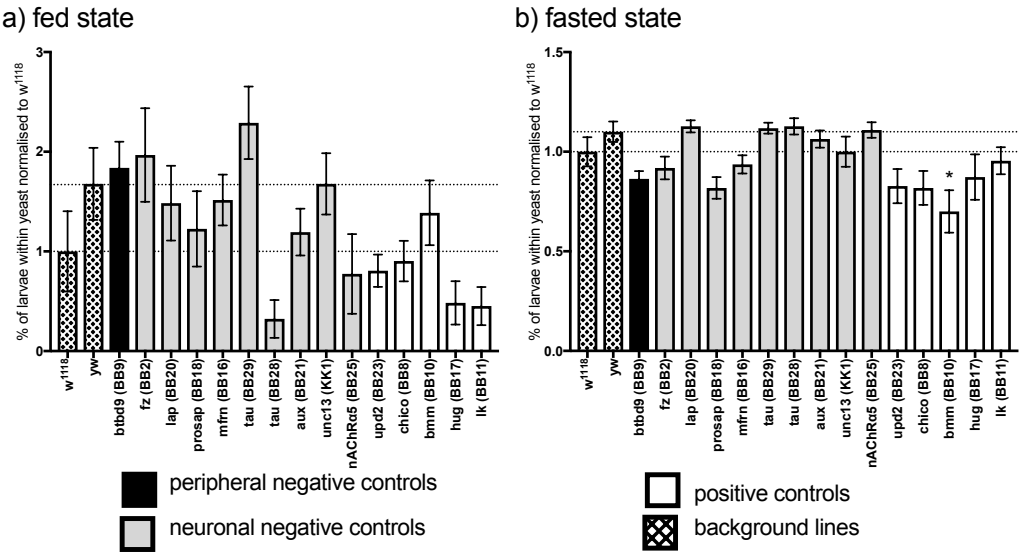


Fig. 3.22 Feeding motivation of *Drosophila* larvae with LoF of control genes in the (a) fed and (b) fasted state. Mean ± SEM is plotted of 7 repeats, each containing 20 larvae. Results were compared to both w^{1118} and yw by Student's t-test. * p<0.05

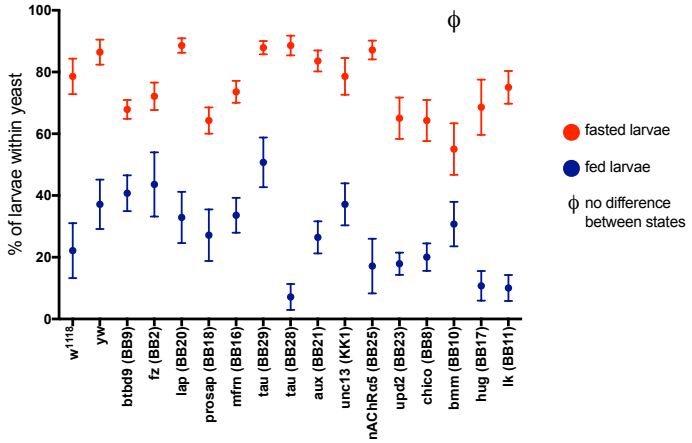


Fig. 3.23 Feeding motivation of *Drosophila* larvae with LoF of control genes in the fed (blue) and fasted (red) state. Mean ± SEM is plotted of 7 repeats, each containing 20 larvae. All genes show a statistically significant difference between fed and fasted, unless marked by ϕ.

3.7.6 Precision of the assays

The results of each assay will be discussed now, rather than in the discussion section of this chapter, because it is important to understand the reasoning for the scoring system developed in section 3.8. This subsection will consider the precision of each assay. Two factors contribute to precision: the sensitivity of the assay (more sensitive assays can detect smaller changes in the measured phenotype) and the reproducibility of the data (how much variation there is between independent repeats).

Among the assays of physiology, the wet mass (figure 3.13), TAG (figure 3.15) and glucose assays (figure 3.16) are precise, as evidenced by their detection of multiple small significant phenotypes. Of these three assays, the results of the control genes suggest that the TAG assay is the most sensitive and the wet mass the least. The starvation resistance assay (figure 3.17) is less precise, but is still able to detect some significant phenotypes. For the dry mass assay (figure 3.14), no significant differences were observed in any control genes which suggests that the assay is insensitive.

As would be expected, the results of the behavioural assays (food intake and movement) show greater variation than the more definitive measurements of physiology. Despite this, the CAFE assay, over-feeding dye assay and dye food absorption assay (figures 3.18, 3.19 and 3.20) were all sensitive enough to detect significant changes in food intake amongst the control genes. The climbing assay (figure 3.21) was not very precise, but was still able to detect a few significant phenotypes.

For the LoF lines, the w^{1118} and yw background flies show a very different phenotype from each other in the over-feeding dye assay (figure 3.19) and the larval feeding assay (figure 3.22). Since experimental lines were compared to both of these backgrounds (section 3.6.1), few significant changes are seen.

3.7.7 Accuracy of the assays

Data from the control genes was also used to assess the accuracy of each assay. This was done by comparing the positive control genes to both the negative control genes, and to published data.

In the screen, lean phenotypes may truly result from the manipulation of a gene important in energy homeostasis. However, any genetic manipulations which cause toxicity and/or non-specific sickness in flies will result in these same phenotypes (Yazdi et al., 2015). Therefore, in all of the assays, phenotypes that were increases compared to background (referred to here as "increase phenotypes") were considered more noteworthy than those that were decreases.

For the wet mass, CAFE and over-feeding dye assays (figures 3.13, 3.18 and 3.19), the positive control genes are enriched for significant increase phenotypes compared to the negative control genes, implying that the results are accurate. This is particularly true of the flies with neuronal RNAi, which is unsurprising given the importance of the brain for feeding behaviour (see chapter 1).

As can be seen in figure 3.22, few lines show a significant phenotype in the larval feeding assay. However, a highly significant difference was detected between fed and fasted larvae for each individual line (figure 3.23) which suggests that the results are accurate. The only line which does not show this behaviour is a mutant in the positive control gene *bmm*, which is consistent with mutations in this gene causing increases in TAG storage (see chapter 1).

For the dry mass assay (figure 3.14), whilst variation was small within each experimental batch, results were very different between batches. This is likely due to environmental factors which influence the drying of the *Drosophila* carcasses such as humidity and temperature. This makes the results inaccurate and so the dry mass assay was not completed for all control lines and was not taken further as part of the screen (preliminary data for chapter 4 genes is in appendix A).

For the dye food absorption assay (figure 3.20), many of the genes that showed an increase in dye absorbance were those which had been noted to have dye on their bodies during collection, meaning that the assay is not an accurate measure of food intake. Therefore, this assay was not completed for all control lines and was not taken further as part of the screen (preliminary data for chapter 4 genes is in appendix A).

The TAG, glucose, starvation resistance and climbing assays (figures 3.15, 3.16, 3.17, and 3.21) do show significant phenotypes in the positive controls, but these genes are not enriched for phenotypes compared to the negative control genes. As explained in section 3.7.2, the

negative control genes are all associated with diseases that either cause or are caused by changes in weight, and so it is possible that they are not true negative controls. Therefore, this lack of discrimination between control genes does not mean that the assays are inaccurate. In particular, the climbing assay is not only a measure of activity, but also of cognitive function and muscle physiology (Chakraborty et al., 2011; Bartholomew et al., 2015), and so the presence of phenotypes in the negative control genes is unsurprising. Therefore, the accuracy of each assay was also assessed by comparing the results obtained here for the positive control genes to published data.

bmm

bmm is a lipase responsible for basal lipolysis in the fat body, the homologue of human *ATGL* (see chapter 1). According to published data, flies with P-element induced KO of *bmm* show almost complete embryonic lethality, but flies lacking only zygotic *bmm* show 100% more TAG as adults than controls and survive 56% longer during starvation (Gronke et al., 2005). These findings were replicated here using three LoF lines with transposon insertions into *bmm*. Firstly, two of the LoF lines were lethal (BB24 and BB27). The viable line (BB10) showed large increases in both TAG and survival during starvation.

chico

chico is a homologue of human IRS1 and is part of the insulin signalling pathway which controls growth and glucose homeostasis in *Drosophila* (see chapter 1). According to published data, adult flies with P-element induced mutation of *chico* have decreased wet mass and dry mass, but increased TAG when normalised to mass, and no change in glucose levels (Bohni et al., 1999; Murillo-Maldonado et al., 2011). Here, two lines were obtained with transposon insertions into *chico*. One of these lines (BB7) was lethal. The other, viable, line (BB8) showed a significant decrease in wet mass and an increase in TAG compared to a w^{1118} background. Although a trend in this direction could be seen when compared to yw, the results did not reach significance. There is also a trend for a decrease in dry mass, but this did not reach significance with either background. Finally, there is a significant decrease in glucose levels which is a different phenotype to that published. Flies with actGAL4-mediated RNAi of *chico* (VV72) were able to replicate both the significantly decreased wet mass and the no change in glucose.

upd2

Published data has shown that flies with deletion of *upd2* have smaller bodies and decreased levels of TAG (Rajan and Perrimon, 2012). Here, a line with a 4.7kb deletion including the *upd2* coding sequence (BB23) was obtained. In contrast to the published results, this line showed significantly increased TAG and no change in mass.

upd1

The *Drosophila* gene *upd1* has been suggested to be a functional homologue of the human gene *leptin*, a hormone with a key role in signalling body TAG stores (see chapter 1). Flies with neuron-specific RNAi of *upd1* have been shown to have increased mass, and increased feeding in a CAFE assay (Beshel et al., 2017). Both of these results were replicated here using flies with neuron-specific RNAi of *upd1* (VV3).

hugin

NmU and its *Drosophila* homologue *hugin* both have important roles in feeding behaviour (see chapter 1). Using tetanus toxin light chain to block synaptic transmission of *hugin*-expressing neurons causes *Drosophila* adults to initiate feeding more quickly upon changing from ND to dye food, although there is no overall change in total food intake (Melcher et al., 2007; Melcher and Pankratz, 2005). Further, RNAi of *hugin* has no effect on larval feeding behaviour (Schoofs et al., 2014). The line used here (BB17) contains an Mi[ET1] element in the *hugin* gene. Like the published data, no change in the dye food absorption assay or larval feeding was seen here. However, no significant difference was found in the over-feeding dye assay.

Leucokinin

The mammalian tachykinins and the fly homologue *Lk* have a role in feeding (see chapter 1). According to published data, flies with disruption of the *Lk* gene or of *Lk*-expressing neurons show increased intake of radioactive/dyed food after starvation, but no changes in the absence of starvation, an increase in starvation resistance, and decreased locomotor activity (Al-Anzi, 2010; Zandawala et al., 2018). Here, flies with a P-element insertion into *Lk* (BB11) also show no change in the dye food absorption assay. For the the over-feeding dye and climbing

assays, there is a significant phenotype in the expected direction when compared to one background control but not the other. No change was seen in starvation resistance.

DSK

CCK is a gastrointestinal hormone involved in the regulation of food intake (see chapter 1). RNAi of the *Drosophila* homologue *DSK* in IPCs or DSK-producing neurons leads to increased ingestion of dyed food following starvation, and increased resistance to starvation (Söderberg et al., 2012). Here, neuronal RNAi of *DSK* (VV36) resulted in an increase in post-starvation food intake, but no change in starvation resistance.

TfAP-2 and twz

TfAP-2 and *twz* function upstream of DSK, and RNAi of each gene in octopaminergic neurons has been shown to increase ingestion of dye food post-starvation and increase food intake via the CAFE assay (Williams et al., 2014). Here, neuronal RNAi was used to knock-down expression of *TfAP2* (VV53) and *twz* (VV54). An increase in both post-starvation food intake and CAFE assay food intake was seen for *TfAP-2*, as expected. No phenotype was seen for *twz*.

Summary

In summary, for all of the assays, the results for at least one of the positive control genes were found to match the previously published phenotypes, as summarised in table 3.4. Since the assay protocols used here were all able to distinguish dietary perturbations (see section 3.4), their ability to replicate the published observations was unsurprising.

Gene	Line	Viable	Wet Mass	Dry Mass	TAG	Glucose	Starvation	CAFE	Over-feeding dye	Dye absorption	Larvae	Climbing
chico	BB7	No	-	-	-	-	-	-	-	-	-	-
	BB8	Yes	trend decrease	trend decrease	trend decrease	decrease	-	-	-	-	-	-
	VV72	Yes	decrease	-	-	no change	-	-	-	-	-	-
bmm	BB24	No	-	-	-	-	-	-	-	-	-	-
	BB27	No	-	-	-	-	-	-	-	-	-	-
	BB10	Yes	-	-	increase	-	increase	-	-	-	-	-
upd2	BB23	Yes	no change	-	increase	-	-	-	-	-	-	-
upd1	VV3	Yes	increase	-	-	-	-	increase	-	-	-	-
hugin	BB17	Yes	-	-	-	-	-	-	no change	no change	no change	-
Lk	BB11	Yes	-	-	-	-	no change	-	increase	no change	-	decrease
DSK	VV36	Yes	-	-	-	-	no change	-	increase	-	-	-
TfAP-2	VV53	Yes	-	-	-	-	-	increase	increase	-	-	-
twz	VV54	Yes	-	-	-	-	-	no change	no change	-	-	-

Table 3.4 Comparison of results obtained to published literature. Blue results agree with the literature, whereas red results do not. A dash indicates that the measurement is not published and/or was not measured here

3.8 Development of a scoring system

The screen produces data for eight phenotypes for each gene: wet mass, TAG levels, glucose levels, starvation resistance, CAFE food intake, dye over-feeding, climbing ability, and larval feeding motivation. To facilitate data analysis, a method of scoring each gene was developed in order to give a measure of the overall phenotype. The algorithm takes a weighted average of the p values for each assay, and subtracts from 1 so that more significant results give higher scores.

The weighting of the assays in the scoring formula was assigned by ranking them in terms of the relevance of the phenotype being measured: the food intake assays were given the highest weighting since this is the phenotype of most interest for this work (human and mouse genetic studies to date have shown that genes which influence energy homeostasis have a strong influence on food intake, see chapter 1), and then the physiology assays came next.

The assays were also ranked according to their precision and accuracy (highest first) based on the results of the positive and negative control genes, as discussed previously. The results of the dry mass and dye food absorption assays are not included, as discussed previously.

Since the summation of 7 phenotypes is 28 (i.e. 7+6+5+4+3+2+1), and of 6 phenotypes is 21, the p value totals are divided by these totals to produce a range of scores with a minimum of 0 and a maximum of 1. Therefore, the final equations for the scoring systems are:

$$Score_{\text{LoFadult}} = 1-(\tfrac{1}{28})(7p_{CAFE}+6p_{dye}+5p_{TAG}+4p_{mass}+3p_{glucose}+2p_{climbing}+p_{starvation})$$

$$Score_{\text{LoFlarvae}} = 1-(\tfrac{1}{2})(p_{larvaefed}+p_{larvaefasted})$$

$$Score_{\text{RNAi}} = 1-(\tfrac{1}{21})(6p_{CAFE}+5p_{dye}+4p_{mass}+3p_{glucose}+2p_{climbing}+p_{starvation})$$

where p_x is the p value in assay x for the gene of interest.

The results of the screen will now be presented according to the *Drosophila* model used (LoF or RNAi) and how well the models were able to distinguish between the negative and positive control genes.

LoF of positive and negative control genes

The phenotypes of flies with LoF of control genes are summarised in figure 3.24. Many of the LoF lines show significant phenotypes in multiple assays and these genes are evenly distributed across the negative and positive control genes. There are more phenotypes which are decreases compared to background than there are increase phenotypes.

The scores are shown in figure 3.25 - for the adult flies there is no significant difference between the positive control genes and the negative control genes. For the larvae, the scores of the positive control genes are significantly increased compared to the neuronal disease negative control genes.

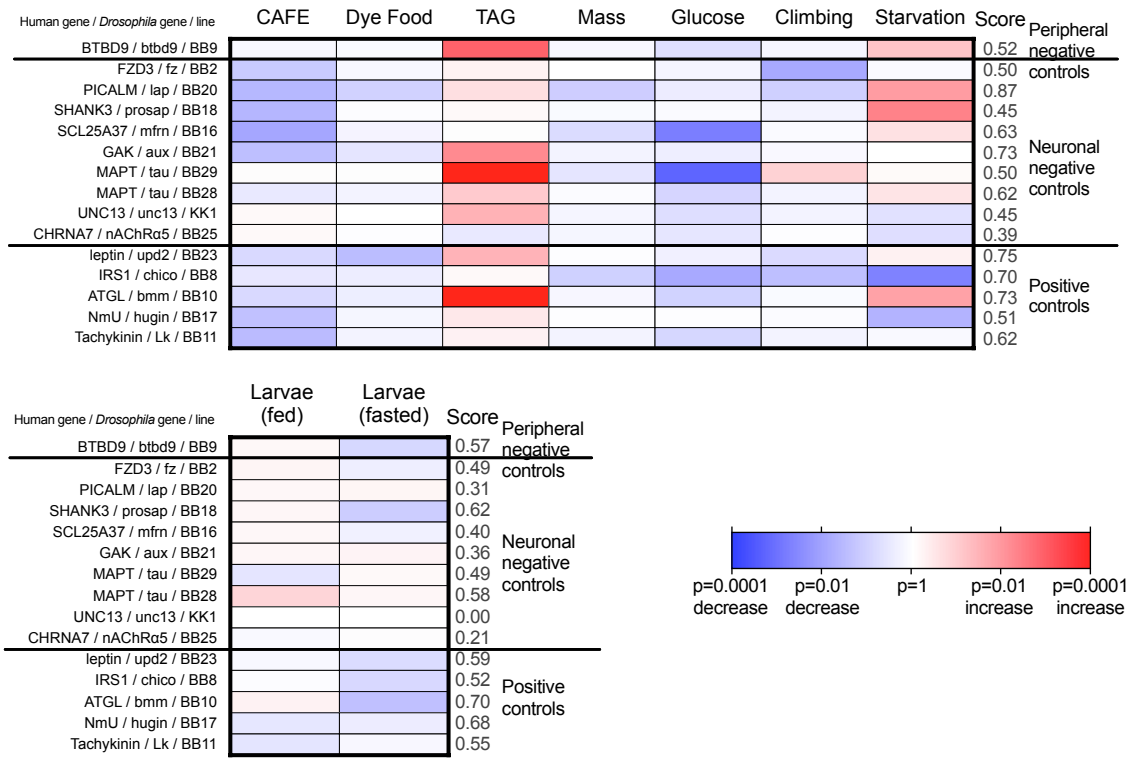


Fig. 3.24 Summary of phenotypes of (top) adult flies and (bottom) larvae with LoF of control genes. Red is an increase in comparison to background and blue is a decrease. * p<0.05, ** p<0.01, ***p<0.001, ****p<0.0001

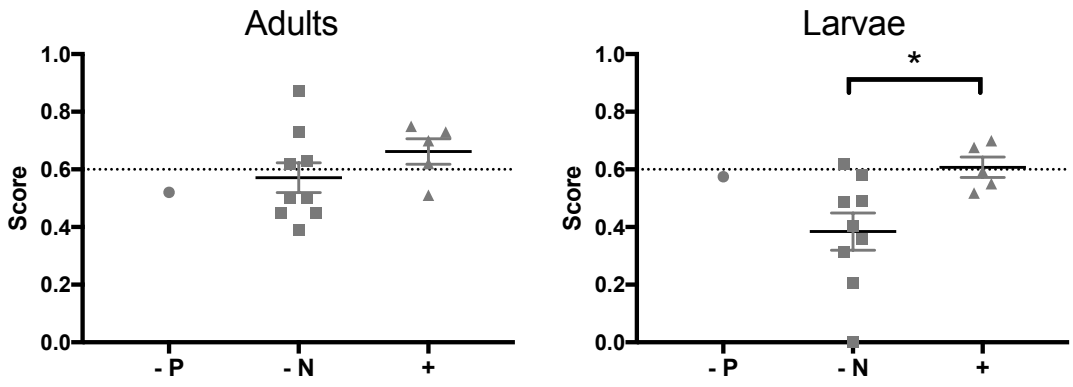


Fig. 3.25 The phenotype score of control genes in LoF *Drosophila*. -P are peripheral disease negative control genes, -N are neuronal disease negative control genes, and + are positive control genes.

Whole body RNAi of positive and negative control genes

The phenotypes of flies with whole body RNAi knock-down of control genes are summarised in figure 3.28. Again, many of the phenotypes are decreases, but the results for most genes are milder (less significant) than when using LoF flies. The positive control genes are enriched for the presence of increase phenotypes compared to the negative control genes. As with the LoF lines, there is a trend towards the positive control genes scoring more highly than the peripheral disease negative control genes, although this is not significant, and there is no difference for the neuronal disease negative control genes (figure 3.27).

After whole body RNAi of *en* (the *Drosophila* orthologue of EN1), the flies were very unhealthy — they were unable to move off of the bottom of the vial, hence their strong decrease phenotype in the climbing assay and the CAFE assay (for which flies have to climb to reach the food capillaries), but a strong increase phenotype in the over-feeding dye assay (they could not move off of the dye food). In agreement with this being an unhealthy phenotype, *en* is known to play a key role in segmentation of the *Drosophila* embryo, and null mutations cause lethality (Kassis, 1990; Gramates et al., 2017). Therefore, since this gene seems to have a non-specific unhealthy phenotype (rather than a specific energy homeostasis phenotype), it should perhaps be removed from the data analysis. In this case, the difference in scores between the positive control genes and the peripheral disease negative control genes approaches significance ($p<0.1$).

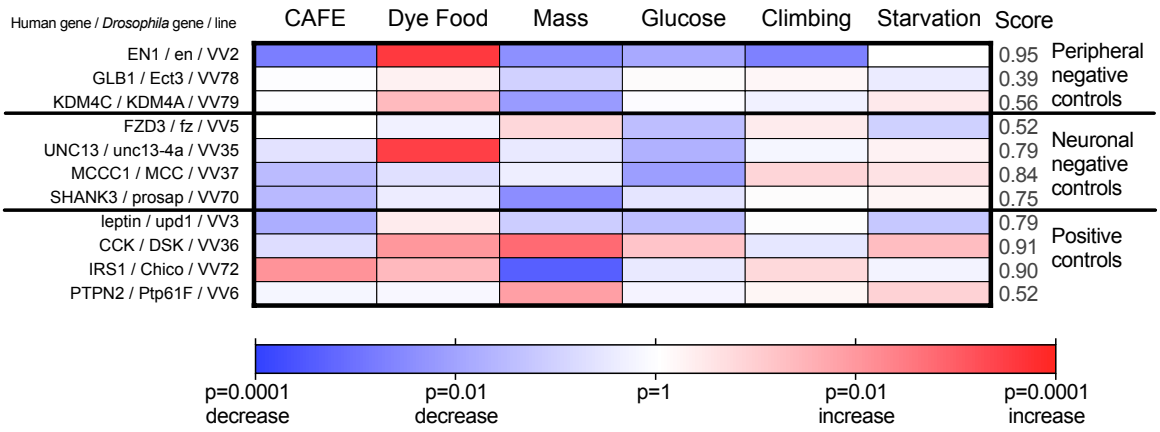


Fig. 3.26 Phenotypes of flies with whole body RNAi of control genes. Red is an increase in comparison to background and blue is a decrease. * p<0.05, ** p<0.01, ***p<0.001

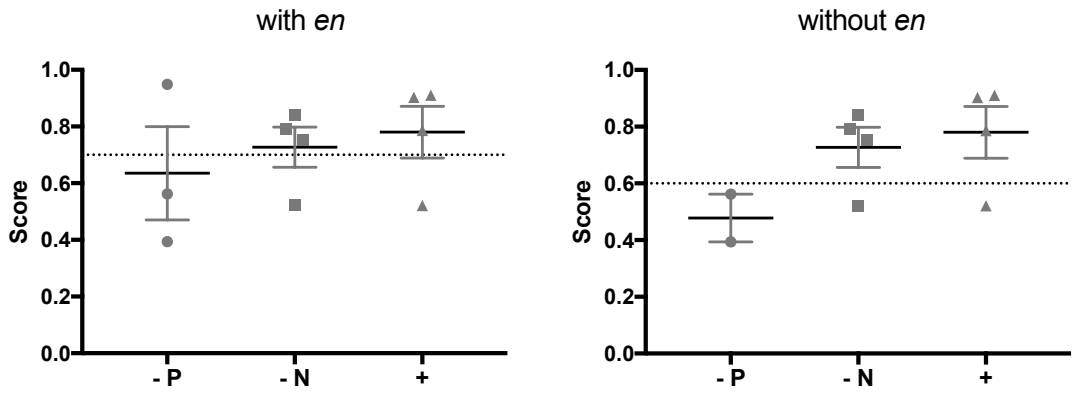


Fig. 3.27 The phenotype score of control genes with whole body RNAi *Drosophila*. -P are peripheral disease negative control genes, -N are neuronal disease negative control genes, and + are positive control genes.

Neuronal RNAi lines of positive and negative control genes

Finally, a summary of the phenotypes of flies with neuronal RNAi of control genes is given in figure 3.28. In this case, few decrease phenotypes are observed, suggesting that neuronal RNAi is generally less detrimental than the whole body genetic perturbations. Further, the positive control genes show many more significant phenotypes than the negative control genes.

Of the neuronal disease negative control genes, only *MCC* (the *Drosophila* orthologue of MCCC1) shows a strong phenotype. MCCC1 was originally included as a neuronal disease negative control because it has been associated by GWAS with Parkinson's disease (Lill, 2016). However, transcriptomics data published in July 2018 showed that this gene is strongly associated with insulin sensitivity and metabolic disease (Timmons et al., 2018). Given this new data, perhaps MCCC1 is not an appropriate gene to use as a negative control. Thus, MCCC1 is not included in any further analysis of the negative control genes.

For the neuronal RNAi, there is a significant distinction between the scores of the positive control genes and both sets of the negative control genes.

There are two low-scoring positive control genes: *chico* and *twz*. It is unsurprising that neuronal RNAi of *chico* does not produce a strong effect, because the reported phenotypes (for example decreased mass, see section 3.7.7) depend on peripheral actions of the gene, not neuronal actions. Therefore, it is perhaps inappropriate to include this gene as a positive control, and removal of this gene increases the significance of the difference between the gene control groups. The reasons that *twz* did not score highly are unclear, and possible reasons will be discussed in section 3.10.

Therefore, neuronal RNAi was able to significantly distinguish between the positive and negative control genes. The threshold for doing so was a score of 0.7.

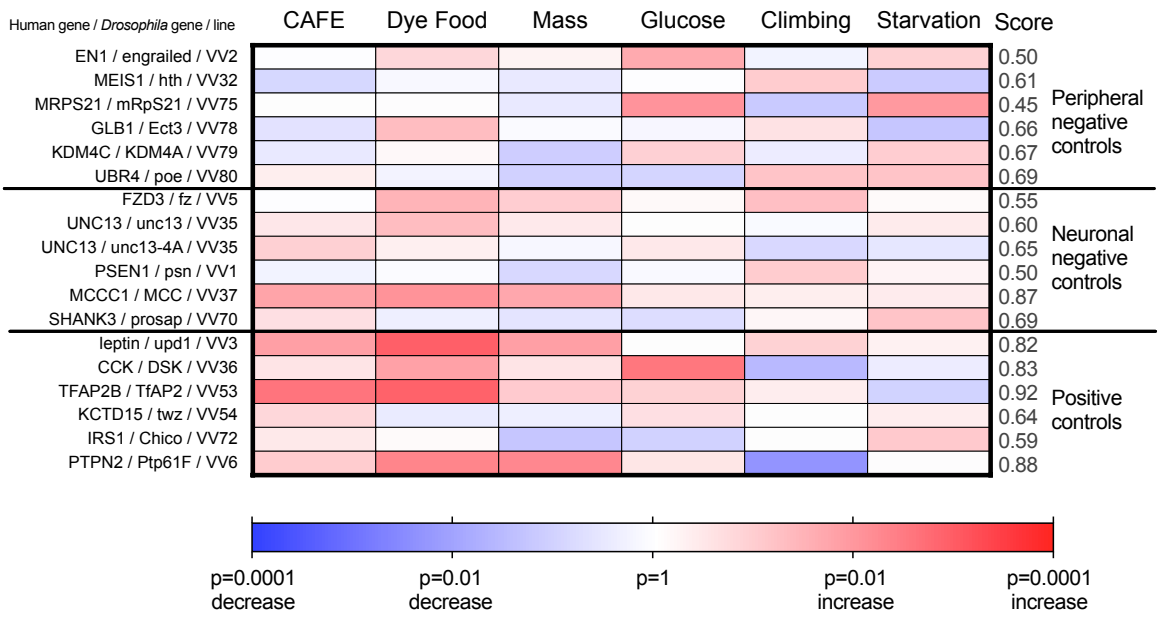


Fig. 3.28 Summary of phenotypes of flies with neuronal RNAi of control genes. Red is an increase in comparison to background control and blue is a decrease. * p<0.05, ** p<0.01

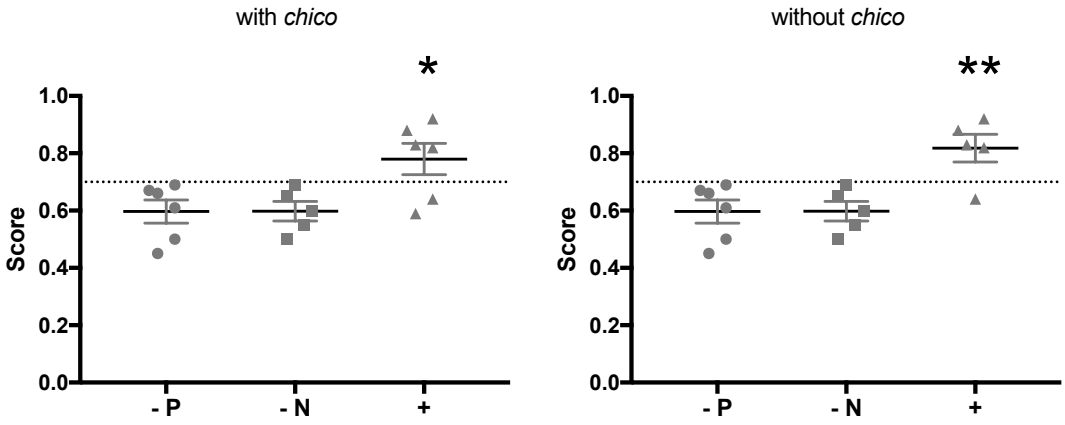


Fig. 3.29 The phenotype score of control genes with whole body RNAi *Drosophila*. -P are peripheral disease negative control genes, -N are neuronal disease negative control genes, and + are positive control genes. *MCC* is not included. * p<0.05.

3.9 Diets

Since high-calorie diets are a major contributory factor to the current obesity epidemic (French et al., 2001), the inclusion of different diets as part of the screen was explored. These diets expose flies to metabolic stress, and thus may reveal additional phenotypes that are not apparent in flies given normal diet (ND) (Smith et al., 2014).

A high-fat diet (HFD) was made by supplementing the ND food with 20% coconut oil, as this has been reported to cause metabolic phenotypes in flies including increased mass, increased TAG storage and decreased lifespan (Birse et al., 2010; Heinrichsen and Haddad, 2012; Ormerod et al., 2017). This same HFD was used to validate the individual screen assays in section 3.4.

The response of the control gene LoF lines to this HFD was tested. Many of the significant results that were found on ND earlier (section 3.7.5) were replicated with the flies exposed to HFD (figure 3.30), which provides evidence that these results are true phenotypes of the lines. Unsurprisingly, the scores (figure 3.31) look similar to those of the ND flies shown earlier (figure 3.25). However, few novel significant increase phenotypes were revealed using HFD.

For some assays (TAG, starvation resistance, and dye over-feeding assays), there is a significant difference between ND- and HFD-fed w^{1118} flies, as shown in figure 3.32. In this case, it is the genes for which no significant change is found between the two diets which are notable, as such flies are insensitive to the dietary change. For example *bmm* does not have increased TAG levels on HFD compared to ND, suggesting that the TAG stores are already maximised on ND, which is consistent with the known role of *bmm* in the break-down of stored TAG.

By comparison, for some assays there is no difference between ND- and HFD-fed w^{1118} flies (figure 3.33, wet mass, glucose, CAFE and climbing assays), in which case it is the flies which *do* show a significant change which are notable, as they have increased sensitivity to the metabolic stress.

However, neither insensitivity nor increased sensitivity to HFD are enriched in the positive control genes compared to the negative control genes.

	Disease	Human Gene	*Drosophila* gene	Wet mass		TAG		Glucose		Starvation		CAFE		Dye food		Climbing	
				ND	HFD	ND	HFD	ND	HFD	ND	HFD	ND	HFD	ND	HFD	ND	HFD
Peripheral negative	RLS	BTBD9	*btbd9*			***	**		*						***		****
	Schizophrenia	FZD3	*frizzled*													*	**
Neuronal negative controls	Alzheimers	PICALM	*lap*							*		*					*
	Autism	SHANK3	*prosap*							**		*					
	Depression	SLC25A37	*mitoferrin*					**	*			*					
	Parkinsons	GAK	*auxilin*			**	*				*	*					
	Parkinsons	MAPT	*tau* (BB29)		**	****	***	***	****		*						
	Parkinsons	MAPT	*tau* (BB28)				**				*						*
	ALS	UNC13A	*unc13*			*			*		*						*
	Multiple	CHRNA7	*nAChRα5*						**								
Positive controls		Leptin	*upd2*											*			***
		IRS1	*chico*		*			*	***	**	****					*	**
		ATGL	*bmm*		*	****	**		**	*	*						
		Neuromedin U	*hugin*						*	*	*	*					
		Tachykinin	*Lk*						***			*					*

Fig. 3.30 Control genes in LoF lines exposed to ND and HFD show similar phenotypes. Red is an increase in comparison to background control and blue is a decrease.
* p<0.05, ** p<0.01, *** p<0.001, ****p<0.0001

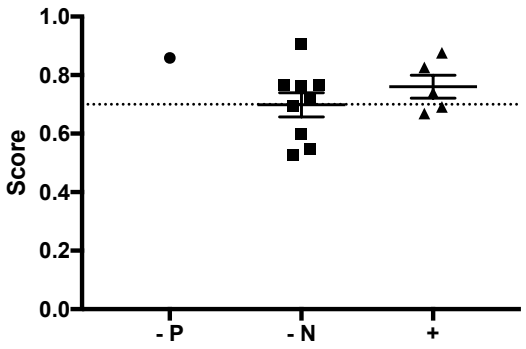


Fig. 3.31 Scores of control genes in LoF lines exposed to HFD

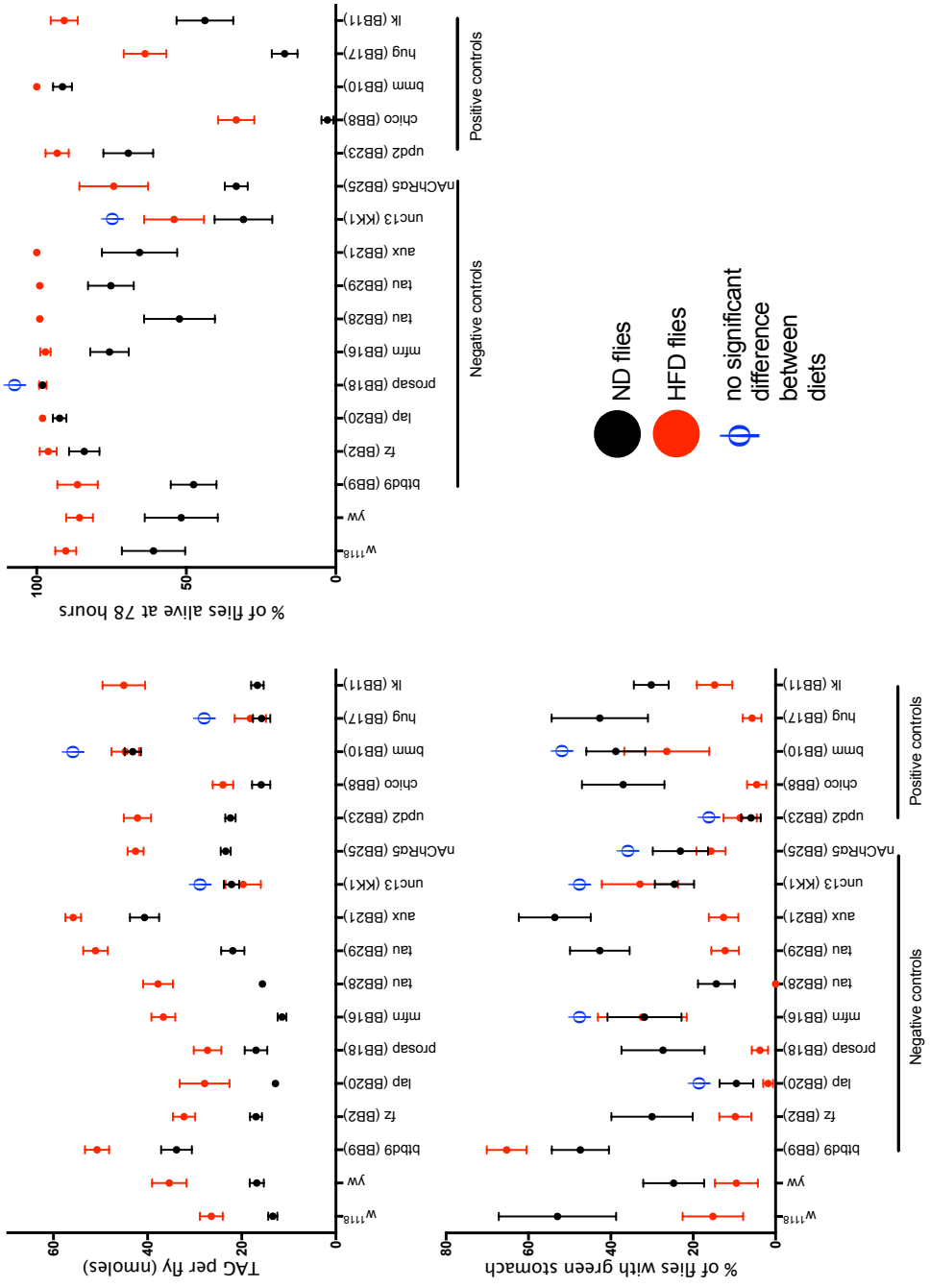


Fig. 3.32 Assays for which w^{1118} and yw flies show a significant difference between ND and HFD. LoF control gene flies which do not show this same phenotype are highlighted with ϕ. Results for ND and HFD were compared by Student's t-test

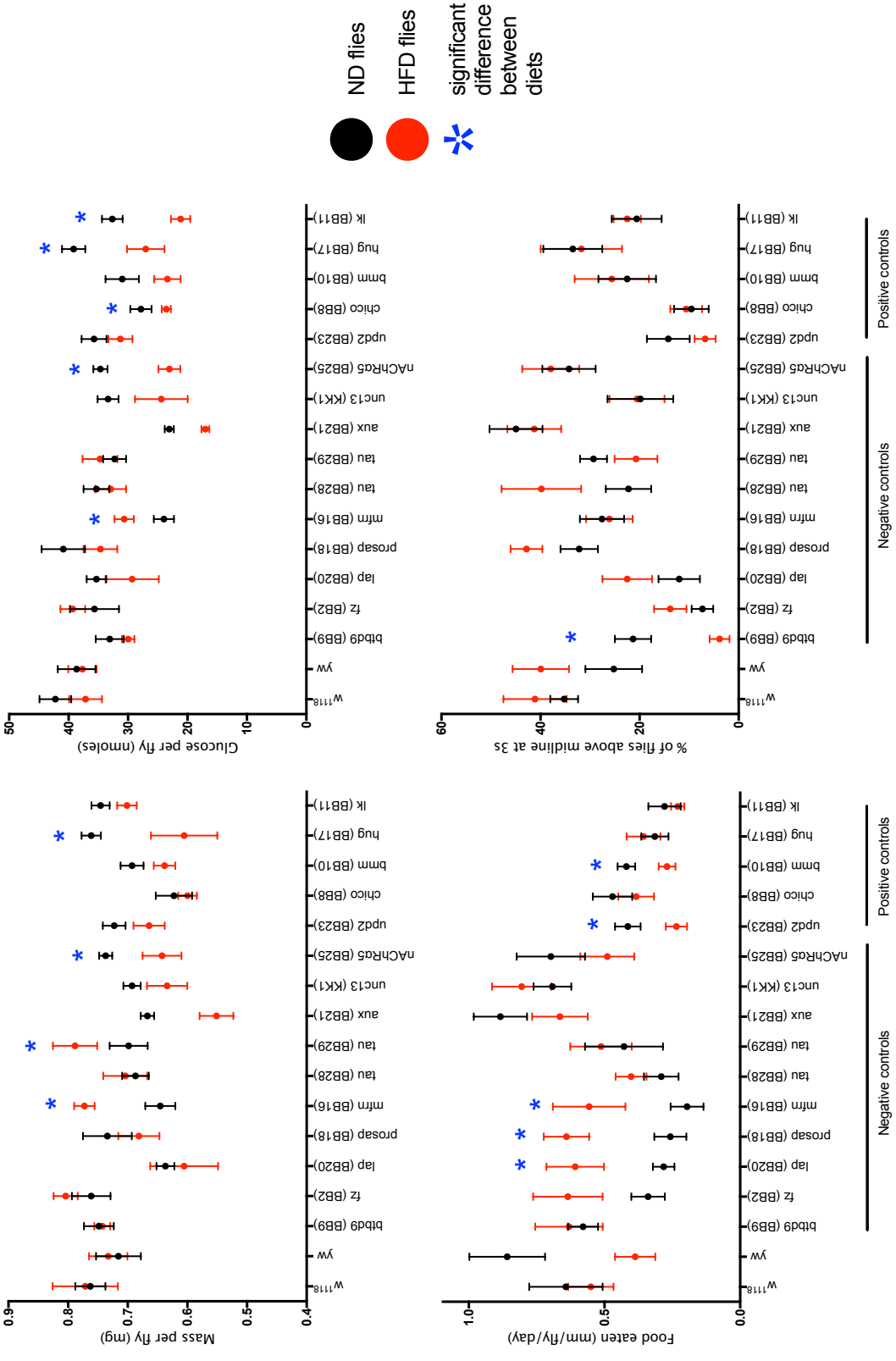


Fig. 3.33 Assays for which w^{1118} and yw flies do not show a significant difference between ND and HFD. LoF control gene flies which do show a phenotype are highlighted with *. Results for ND and HFD were compared by Student's t-test

The HFD had to be kept below 20°C to prevent the coconut oil melting, but GAL4 is less active at this temperature (Busson and Pret, 2007) and so HFD was not suitable for the RNAi flies. Instead, a high sugar diet (HSD) was tested containing 35% sucrose and glucose (compared to 15% in ND) which has been reported to cause strong metabolic phenotypes in flies (Morris et al., 2012b; Musselman et al., 2011; Na et al., 2013; Pasco and Leopold, 2012; Tennessen et al., 2014; Trindade de Paula et al., 2016; May et al., 2015). The effects of this diet were tested with w^{1118} flies. As can be seen in figure 3.34, no difference between ND- and HSD-fed flies was seen in most assays. Therefore HSD was not continued for further use as part of the screen (some preliminary data is in appendix A).

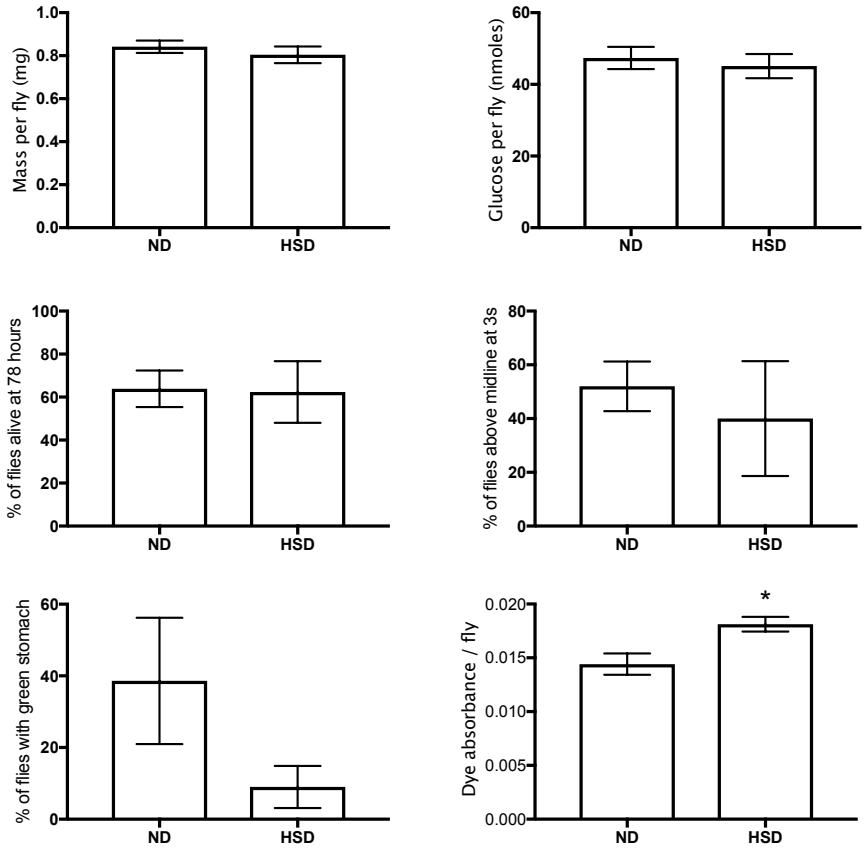


Fig. 3.34 Phenotype of w^{1118} flies exposed to five days of either ND or HSD.
Mean ± SEM is plotted. Results for ND and HSD were compared by Student's t-test.

3.10 Discussion

There is an unmet need for a high-throughput method to investigate human genetic data related to the regulation of energy homeostasis. The aim of this chapter was to create and validate a high-throughput screen of energy homeostasis phenotypes using *Drosophila* as a model to study specific human genes of interest.

3.10.1 *Drosophila* models and their backgrounds

Several large collections of *Drosophila* stocks for gene manipulation are readily available from public stock centres. Two models were used in this chapter — LoF and RNAi — and the basic principles underlying their genetic perturbations were outlined in section 3.2. The two models required use of different background lines for statistical comparison.

For many LoF lines, the background is either not known or not available for use. Further, lines obtained from stock centres are inbred, which is known to cause changes in gene expression, disproportionately affecting genes involved in metabolism (Kristensen, 2005). According to the Bloomington stock centre (from which most of the LoF lines were obtained): "For many collections there likely isn't a good control". Backcrossing two LoF lines to w^{1118} showed that background was an important factor for the assays of energy homeostasis, with strong phenotypes in inbred flies disappearing after backcrossing (section 3.6.1). Therefore, the use of matching backgrounds for comparison for each line is important. However, backcrossing stocks reduces the speed of the screen, can only be applied to lines where the mutation has a visible marker, and can introduce novel mutations, so is not practical for use in a high-throughput screen. Instead, the variability and uncertainty in the background of the LoF lines was overcome by using multiple backgrounds for statistical comparison (yw and w^{1118}).

For RNAi, the backgrounds of the lines were known (KK-GAL4 or GD-GAL4) and easily obtainable for phenotyping (section 3.6.2). There are two exceptions to confident use of these backgrounds. Firstly, the compound X stocks may have slightly altered backgrounds due to the crosses required to remove the compound X balancer chromosome. Secondly, some of the KK siRNA lines cause over-expression of the *tiptop* gene which produces a strong phenotype, and thus requires use of an additional background line.

3.10.2 Assay selection

Ten assays were found in the literature that, when combined together to form a screen, allow measurement of a number of characteristics related to energy homeostasis in *Drosophila*: five assays that measure physiology, four assays that measure food intake, and one that measures movement (section 3.3). These assays were all found to be high-throughput enough for use in a large scale screen.

In addition, an assay was created that is able to detect feeding motivation in adult flies (section 3.5). This assay was quite low-throughput and thus was intended for use on just a few genes which the high-throughput assays indicate are top hits.

Several high-throughput genetic screens of metabolic phenotypes have been published, but these generally use just one assay, for example TAG (Baranski et al., 2018) or glucose (Ugrankar et al., 2015). To my knowledge, there are no published results in which a suite of assays was used to screen a large number of genes for energy homeostasis phenotypes. Using multiple assays in a screen has two advantages. Firstly, the screen is less likely to give false positive results because anomalous results would have to be obtained for several assays. Secondly, genes which do not show a large phenotype in any single assay, but which have subtle changes in multiple assays will also be highlighted, making the screen more sensitive. This sensitivity is particularly important given that behavioural traits (such as feeding) are difficult to measure precisely. The disadvantage of using a multi-assay screen is that results generation is slower.

3.10.3 Assay validation via metabolic perturbation

The high-throughput assays were all validated for use by their ability to differentiate w^{1118} flies given ND from flies with metabolic perturbation, either given HFD or fasted (section 3.4). Significant differences were detected in the expected direction for all assays between ND and fasted flies. The phenotype of the HFD-fed flies is less strong, likely because these flies were kept in different conditions in order to stop the HFD food melting. These conditions (lower temperature, lower humidity, and no light) have been shown to affect metabolic phenotypes in *Drosophila* (Al-Saffar et al., 1996; Caravaca and Lei, 2016) which complicates interpretation of the results.

Based solely on the metabolic perturbation validation, the TAG and CAFE assays seemed the most sensitive, able to detect both dietary perturbations, whereas the dye food absorption and dry mass assays were potentially inaccurate.

The TAG assay could not be done using the RNAi lines as the levels of TAG in the sample were always below the detection limit. This is unlikely to be due to user error or contamination of reagents because the samples were extracted and analysed at the same time as samples from LoF lines which did show detectable levels of TAG. It is also unlikely to be an anomalous result, since >300 RNAi samples were tested in this manner. This suggests that perhaps the background of the RNAi lines contain low levels of TAG. This is unfortunate because, based on the results from the metabolic validation, control genes validation and HFD trial experiments, the TAG assay seemed to be one of the most accurate and sensitive assays.

It is possible to do the larval feeding assay on RNAi lines, but the requirement to collect such a large number of UAS_G-siRNA virgins means that this assay would not be high-throughput.

All other assays could be used in both models.

3.10.4 Validation using control genes

Identification of control genes

Previous *Drosophila* screens of GWAS loci, for example by Baranski et al. (2018) and Pendse et al. (2013), have used just one assay to probe their genes of interest, and so were able to compare their results to published genome-wide screens of the same assay. Both papers found that their phenotype of interest was enriched in their GWAS genes compared to the genome average. However, such an approach was not possible for the suite of assays used here, since there is not, to my knowledge, any published work for a genome-wide screen using most of the assays.

Instead, the *Drosophila*-based screen for energy homeostasis phenotypes was tested using a series of negative and positive control genes (section 3.7). This allowed the screen to be validated by its ability to differentiate between these two types of controls, and the difference to be quantified to provide a threshold for interpretation of experimental genes in later chapters. The positive and negative control genes were phenotyped in a blinded manner to avoid biasing of the results.

There were some differences in the way in which the positive and negative control genes were selected. Firstly, most of the positive control genes were identified from published *Drosophila* literature, whereas the negative control genes did not have any similar data published. Further, the positive control genes have proven functional homology between

humans and flies, but the orthologues of the negative control genes were identified using only sequence homology, which does not guarantee conservation of function. Therefore, whilst the results from the control genes can be compared, there is a limit on the confidence associated with the comparison.

18 genes were chosen as negative controls. Whether these genes are truly suitable for this purpose is not certain, given that they are all associated with diseases that are, in some way, related to body weight/fat (the neuronal control diseases have a weight-related symptom, and the risk of developing the peripheral control diseases is affected by weight). This is exemplified by the recent data showing that MCCC1, which was chosen as a negative control gene due to its association with Parkinson's disease (Lill, 2016), is also associated with metabolic disease (Timmons et al., 2018). However, since the purpose of the screen is candidate gene selection, it is important to avoid false positive results, and so setting the threshold at a high level by using these negative controls is perhaps advantageous.

The negative control genes were chosen due to their association with human disease. In order to prevent bias in gene selection, published *Drosophila* phenotypes of these genes were not researched. However, after the phenotyping data was complete, a literature search revealed that *en* plays a key role in segmentation of the *Drosophila* embryo (Kassis, 1990; Gramates et al., 2017). Therefore, the unhealthy phenotype observed is unsurprising, and this gene was (retrospectively) considered unsuitable for use as a negative control.

According to published microarray data, all of the chosen control genes are expressed in the larval and/or adult CNS in *Drosophila*, which is an important consideration for the neuronal RNAi data.

Comparison of positive control genes to published data

The accuracy of the assays was partially assessed by comparing the results of the positive control genes to published data. For some genes, the protocols used here were able to replicate all of the published phenotypes (e.g. *upd1*, *bmm* and *TfAP-2*). Most of the other genes were able to replicate at least some of the data. For those cases in which no phenotype or the opposite phenotype was seen in an assay, this is likely due to differences in experimental paradigms:

- The assay protocols used are not identical. For example, Söderberg *et al.* saw an increase in starvation resistance for *DSK* after single-housing their flies (Söderberg et al., 2012), but here no change was seen here where the flies were kept in groups of 15.

- The genetic perturbation was different. For example, Melcher *et al.* saw an effect on feeding after blocking transmission of *hugin*-expressing neurons using tetanus toxin light chain (Melcher and Pankratz, 2005; Melcher et al., 2006, 2007), whereas no effect was seen here using RNAi against *hugin* itself.
- There is variation in the effectiveness of different genetic perturbations of the same gene: the precise location of a P-element integration in LoF lines affects how much the target gene is perturbed (Groth et al., 2004), and siRNAs are differentially expressed depending on their genomic location (Heigwer et al., 2018). For example, two of the *bmm* LoF lines were lethal, but one was not, despite all three lines having a transposable element inserted into the gene. This difference in viability between lines which affect the same gene suggests that there is also variability in energy homeostasis phenotypes. Generally, the *Drosophila* lines used in the literature were not identical to those used here.
- The GAL4 lines used with the RNAi flies had a different expression pattern. For example, Söderberg *et al.* performed RNAi in DSK-producing neurons only, whereas here, elav-GAL4 was used to perform pan-neuronal RNAi, and the starvation resistance phenotype was not seen.
- One or both of the lines used as background for the LoF lines was not appropriate. For example, flies with *chico* LoF showed decreased glucose, but might not have had they been compared to their actual background of y^1w^{67c23}.
- The assay is inaccurate and/or insensitive, for example the dry mass assay (section 3.7.7).

There is only one gene for which none of these factors can explain the lack of phenotypes observed: *twz*. The same *twz* siRNA line was used as in the literature (Williams et al., 2014), discounting siRNA variability as the cause of the absent phenotype. There were multiple differences in the methods used (different food, temperature, GAL4 driver), but these same arguments also apply to *TfAP-2* (from the same paper) for which a phenotype was observed here, so these cannot explain the lack of phenotype. Thus, the reason for this lack of phenotype seen with *twz* remain unclear.

Precision and accuracy of the assays

Phenotyping of the positive and negative control lines allowed the assays to be tested for their use in a large-scale screen. The precision and accuracy of each assay were discussed

in detail earlier (sections 3.7.6 and 3.7.7), and so will not be discussed again, just briefly summarised in the following table. The precision refers to both the sensitivity of the assay and the variability of the phenotype, as evidenced by how small a change was detected as significant. The accuracy refers to whether the positive control genes were enriched for significant phenotypes compared to the negative control genes, and whether published data could be replicated with the positive controls.

Assay	Precision	Accuracy	Included in screen
CAFE assay	High	High	Yes
Over-feeding dye	Medium	High	Yes
TAG	High	High	Yes
Wet mass	High	High	Yes
Glucose	High	High	Yes
Climbing	Low	Medium	Yes
Starvation resistance	Low	High	Yes
Larvae	Low	High	Yes
Dye food absorption	High	Low	No
Dry mass	Low	Uncertain	No

The assays were all considered suitable for further use, except the dry mass and dye absorption assays which were found to be inaccurate and/or imprecise.

Scoring system

Based on the results of the positive and negative control genes, a scoring system was developed using the results for each assay (section 3.8). An advantage of this scoring system is that genes which show subtle phenotypes in multiple assays will score more highly than genes which show strong phenotypes in just one assay, which both highlights genes that would not be detected by a screen consisting of only one assay, and also makes false positive results less likely. There are two limitations of the scoring system. Firstly, results which are increases compared to background are not distinguished from those that are decreases, and thus flies which are unhealthy (rather than having a specific energy homeostasis phenotype) may also score highly, as was seen with *en*-actGAL4. Thus, the scores should not be viewed in isolation from the assay data. Secondly, as discussed earlier, the negative control genes may not be totally free of energy homeostasis phenotypes, and so choosing a cut-off score of 0.7 based on these genes limits the sensitivity of the screen.

Control genes in different *Drosophila* models

In both the negative and positive control genes, many of the LoF lines show significant phenotypes in multiple assays. Since many of these phenotypes are maintained across ND and HFD, this implies that the results do accurately represent the phenotype of the lines. However, the lines may not truly represent the phenotype of the genes for two reasons. Firstly, the background lines used for statistical comparison (w^{1118} and yw) may not be appropriate for every line, generating false positive/negative results. Secondly, most of the significant results are decreases (compared to background) rather than increases, and although this leanness may be a true phenotype, it may just be non-specific unhealthiness. The scoring system found a trend towards the positive control genes scoring more highly than the peripheral disease negative control genes, but it was not significant.

Similar to the LoF flies, flies with whole body RNAi showed many decrease phenotypes. However, the results are milder than LoF, likely because RNAi does not completely stop gene expression. Some increase phenotypes can also be seen, and these are enriched in the positive control genes. Again, a trend is seen in the scores, but it does not reach significance.

Finally, for neuronal RNAi, few decrease phenotypes are observed, suggesting that the neuronal RNAi is not commonly detrimental. Further, the positive control genes show many more increase phenotypes than the negative control genes. There was a significant difference between the positive control genes and the negative control genes.

3.10.5 Use of diets

Since the aim of studying obesity genetics is to understand common obesity, and modern high-calorie diets are thought to be a major factor in the current obesity epidemic, the use of a HFD and a HSD as part of the *Drosophila* screen was explored (section 3.9). Both diets have been reported to cause metabolic phenotypes in flies. However, when tested here, the HFD did not reveal any genes which showed a different response to HFD across multiple assays, and the HSD did not cause any detectable phenotype. This is likely because most of the published papers exposed flies to the diet for several weeks (as opposed to the five days used here), and many used female flies (as opposed to males, which feed less than females and so will be less affected by the change in diet). Further, the published literature has conflicting evidence for the effect of these diets on the phenotypes of interest. For example, HSD has been shown to increase (Morris et al., 2012b), decrease (Musselman et al., 2011)

and not change (May et al., 2015) the wet mass of *Drosophila*. Therefore, neither diet were considered further for inclusion as part of the screen.

3.10.6 Comparison of the LoF and RNAi models

The two *Drosophila* models used to assay the positive and negative control genes - LoF and RNAi - will now be compared for their suitability for the high-throughput screen.

Background: All of the RNAi lines came from a known background (KK or GD) which are available to use for comparison. For >50% of the LoF lines, the background was not known, and the rest came from a range of backgrounds, which has a strong effect on phenotype (section 3.6.1).

Assays: All of the assays can be applied to LoF lines (but only if adults are viable). However, the TAG assay, larval assay and use of HFD could not be performed on RNAi flies (sections 3.4 and 3.9).

Viability: 25% of *Drosophila* genes are lethal when mutant (Wangler et al., 2017). This figure is slightly lower for whole body RNAi because some residual gene expression remains, and even lower for neuronal RNAi (for example, see chapter 4).

Results: RNAi was able to significantly distinguish between positive and negative control genes, whereas LoF was not (section 3.8).

Speed: It was found that the LoF and RNAi models produced results at approximately the same rate. However, the LoF lines can be assayed as soon as the first generation are born, whereas phenotyping of RNAi flies is delayed by two weeks because the siRNA lines must be crossed to GAL4 and then the offspring assayed. Further, some siRNA lines require removal of a compound X balancer chromosome, which takes three generations.

Tissue-specificity: LoF lines have universal disruption of the gene of interest, which makes many of the lines unhealthy, thus potentially masking energy homeostasis phenotypes. By contrast, RNAi is tissue-specific, which reduces the lethality and sickness associated with universal mutation, whilst also providing some mechanistic information (location of gene action) about the gene being studied.

Efficiency: The efficiency of both LoF and RNAi is known to be variable. For LoF, the exact position of the P-element integration influences gene expression (Groth et al., 2004). For RNAi, the design of the siRNA and the location of genomic integration both affect the efficiency of knock-down (Dietzl et al., 2007). With LoF it is possible to completely stop expression of the gene of interest, whereas with RNAi the knock-down can only be partial

(generally to a level of 25%, Heigwer et al. (2018)) which may not be enough to affect some genes.

Genetic specificity: Both LoF and RNAi can have off-target effects. For RNAi, this happens when an siRNA has sequence similarity to non-target RNAs (Seinen et al., 2011). For LoF stocks this can happen if the line contains a deletion of a genomic locus encompassing several genes, or if the location of the P-element influences multiple genes (Groth et al., 2004).

Availability: For both models, mutant flies can be sourced from publicly available stock centres. Approximately 70% of *Drosophila* genes have lines available with P-element insertions (Hummel and Klämbt, 2008), whereas 91% of genes have siRNA lines available. LoF lines are generally cheaper than RNAi lines.

On the basis of all of these factors, it was decided that RNAi was the more suitable model for use in a high-throughput screen of genes for effects on energy homeostasis and food intake.

3.11 Summary

In this chapter, a collection of assays were identified for investigation of physiology, energy intake and movement. They were validated by their ability to detect dietary perturbation of wild type flies. Two genetic models (LoF and RNAi) were explored in *Drosophila* for their use in the screen. Next, a series of positive and negative control genes were identified and assayed in order to validate the screen by showing that it is able to differentiate between the controls, and RNAi was found to be a better model than LoF for doing so. Thus, the framework of a *Drosophila*-based genetic screen for energy homeostasis phenotypes was established.

Chapter 4

A *Drosophila*-based screen of genes identified by GWAS to influence BMI

4.1 GWAS and polygenic obesity

As discussed in chapter 1, genetics is a powerful tool with which to identify and study the pathways involved in energy homeostasis in humans. In the past, rare monogenic and syndromic obesity disorders have provided invaluable insights into the control of body weight and food intake. Now, advances in technology and the accessibility of genomic resources covering large general populations is beginning to allow the relationship between genes and common obesity to be investigated. It is hoped that these studies of complex polygenic obesity will bring new discoveries that advance our understanding of energy homeostasis, and eventually facilitate development of new therapies for obesity. To date, multiple GWASs (see chapter 1) have been performed on general populations for various obesity-related traits such as BMI, and these have uncovered many novel genetic loci which are the subject of this chapter.

4.2 Moving from GWAS to biological understanding

BMI is an approximate measure of body fat and obesity. Between 2007 and 2013, GWASs of hundreds of thousands of people had identified 36 genetic loci which are robustly associated with BMI (Speliotes et al., 2010; Lu and Loos, 2013). Some of these loci are near genes that were already known to have a role in energy homeostasis, for example MC4R and POMC. But for many of the genetic loci discovered by GWAS, very little is known about the biology of the nearby genes.

To move from the genetic findings of GWAS to development of therapies, the genes responsible for the GWAS signals must be studied. However, the speed of converting these GWAS statistical associations into an understanding of the underlying biological mechanisms has been slow. Since 2013, 168 novel genetic loci associated with BMI have been identified by GWAS (Locke et al., 2015; Akiyama et al., 2017). But in the same time period, only a handful of GWAS BMI genes have had functional studies published in mouse models: FTO (Speakman, 2015; Tung et al., 2015), TMEM18 (Larder et al., 2017), NEGR1 (Lee et al., 2012) and CADM2 (Rathjen et al., 2017). There are three reasons underlying this slow speed of progress.

Firstly, GWASs generate large lists of data - to date 204 genetic loci have been associated with BMI. This problem is compounded by the fact that GWASs lack the resolution to specifically identify the causal gene. As was explained in chapter 1, GWASs genotype millions of SNPs across the entire genome. These SNPs represent the common variation in their region of the genome and exist in linkage disequilibrium with several hundred kilobases containing multiple genes and regulatory elements of remote genes (Nicolae et al., 2010; Wall and Pritchard, 2003; Wangler et al., 2017). This means that it is not clear from the GWAS data alone which gene is causing the effect on energy homeostasis, thereby increasing the number of genes which must be functionally investigated.

Secondly, rodent models have been invaluable in studying monogenic obesity and to date they have been the model organism of choice for studying the GWAS genes. But as discussed in chapter 1, experiments with mice are low-throughput and expensive. Unfortunately the scientific community lacks the resources to study all of the GWAS genes in detail in mice.

Finally, use of gene expression microarrays has shown that gene expression is enriched in the CNS for the GWAS BMI loci genes (Locke et al., 2015). However, the difficulty of obtaining tissue samples from the brain makes it a complicated system to study.

Currently the genetic variation that has been identified by GWAS only accounts for less than 10% of the heritability of body weight (Bogardus, 2009; Speakman et al., 2018). Therefore it is likely that more genetic loci will be found, and improvements in technology mean that they will be found at an increasing rate.

This discrepancy in speed between the identification and the validation of genes calls for use of a high-throughput model organism to screen the GWAS candidate genes, so allowing resources to be focussed on studying only the most promising genes in mammalian models.

4.3 *Drosophila* as a tool to investigate GWAS loci

High-throughput *in vivo* screens using *Drosophila* have the potential to assess candidate genes from GWAS in an unbiased manner, and so accelerate functional studies of GWAS data. As exemplars, two recent papers have used *Drosophila* to functionally evaluate genes from GWASs of diseases of energy homeostasis.

Pendse et al. (2013) studied 38 human genetic loci that GWAS had found to be associated with type 2 diabetes. They took the 130 human genes which lie within an 100kb window of each of the GWAS SNPs, and identified fly orthologues for 71 of these genes. Using RNAi to knockdown each gene in the whole body, Pendse et al. tested the flies for sucrose-dependent toxicity. Most (33/38) of the GWAS SNPs were able to be studied in this manner. 42% of the *Drosophila* genes showed a sucrose-toxicity phenotype, which is much higher than the genome average of 9% which affect glucose levels (Ugrankar et al., 2015). At 13 loci, only 1 gene showed a phenotype in flies, and at 9 loci multiple human genes had a fly phenotype. Some of these genes have proven effects on phenotypes of relevance to diabetes in mouse studies. Based on their results Pendse et al. suggested which novel genes should be the focus of further studies in humans/mice.

Very recently, Baranski et al. (2018) used *Drosophila* to examine 78 genetic loci which GWASs have associated with BMI. They looked at all genes within a 250kb radius of each SNP, and found at least one gene had a fly orthologue at 62 of these loci. After RNAi knockdown of each gene in the brain and fat body, they measured the amount of TAG in the flies. Of the 62 testable loci, 26 (42%) contained at least one gene with a significant TAG phenotype, which is higher than the genome-wide average of 5% (Pospisilik et al., 2010). Further, at 10 of these 26 loci, the gene producing a phenotype was not the gene closest to the SNP, which led the authors to suggest that further studies in rodents or humans should focus on these more distant genes, which would perhaps not be obvious just from looking at the GWAS statistical data.

In both of these studies, the measured phenotypes are considerably enriched within the genes from the diabetes/BMI GWAS loci compared to the genome average. This suggests that there are high levels of conservation of energy homeostasis pathways during evolution, and provides evidence for the utility of a *Drosophila*-based high-throughput functional genetic screening strategy.

4.4 Chapter aims

GWASs have provided a glut of potential insights into the genetics underlying obesity, but there remains the need to study the biological mechanisms of these statistical associations.

This chapter focusses on genetic loci which GWASs have found to be associated with human BMI. This statistical data was functionally investigated using the high-throughput *Drosophila*-based screen for energy homeostasis phenotypes which was developed in chapter 3. This data set was chosen because some of the genes have already been functionally explored, so allowing verification of the screen results. But most of the genes have not yet been studied, thus also allowing for the potential discovery of novel results. Further, mouse studies of a few of the GWAS-identified genes within the lab, for example TMEM18 (Larder et al., 2017), have shown that their effects on energy homeostasis phenotypes are relatively small, and so if the *Drosophila*-based screen is able to detect such genes (proof-of-concept), then it is likely to also be sensitive enough to use on other types of genetic data.

Although similar to the work of Baranski et al. (2018), this work used a suite of assays to evaluate multiple phenotypes including feeding behaviour, rather than only examining TAG levels.

4.5 Identification of human and *Drosophila* genes for study

4.5.1 Selection of human genes for study

The work in section 4.5.1 was done by Dr Loraine Tung. Briefly, by 2014, published GWASs had identified 36 loci robustly associated with BMI (Speliotes et al., 2010; Lu and Loos, 2013), as shown in figure 4.1.

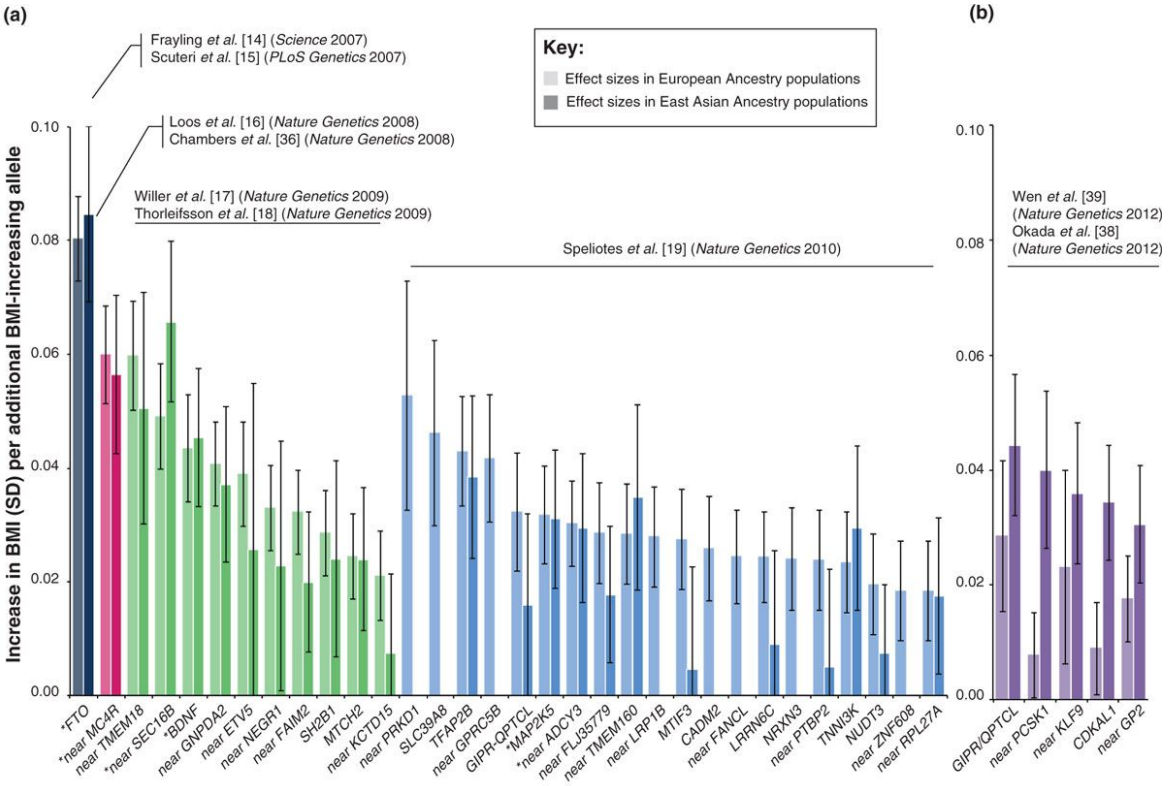


Fig. 4.1 Figure from Lu and Loos (2013) showing the GWAS loci that are associated with BMI and were investigated in this thesis. SNPs are annotated by the nearest gene.

At 28 of these loci, just the gene closest to the SNP was selected for study: FTO, MC4R, BDNF, TMEM160, FANCL, CADM2, PRKD1, LRP1B, PTBP2, MTIF3, ZNF608, RPL27A, NUDT3, PCSK1, KLF9, GP2, CDKAL1, LRRN6C, TNNI3K, SLC39A8, ETV5, MTCH2, KCTD15, SEC16B, TFAP2B, FAIM2, NRXN3, GPRC5B.

At the remaining 8 loci, multiple genes were studied in order to investigate whether the screen could be used to differentiate between multiple genes at the same locus. At each of the following 6 loci, genes whose expression level correlates with the SNP variant (Speliotes et al., 2010) were studied in addition to the nearest (underlined) gene:

- The rs7359397 locus: SH2B1, APOB48R, SULT1A2, AC138894.2, ATXN2L, SBK1 and TUFM.
- The rs713586 locus: RBJ, ADCY3, and POMC
- The rs2241423 locus: MAP2K5 and LBXCOR1
- The rs2287019 locus: QPCTL and GIPR
- The rs2112347 locus: FLJ35779, HMGCR, COL4A3BP
- The rs10938397 locus: GNPDA2 and GABRG1

In addition, 2 SNPs were screened in a more systematic manner. These were chosen because the lab group are studying mouse models of genes present within these loci:

- The rs2815752 locus: the closest gene is NEGR1. In addition, ZRANB2, PTGER3 and LRRIQ3 lie within 1500kb as well as a binding site for the transcription factor NKX6.1 (Wheeler et al., 2013).
- The rs2867125 locus, at which all genes within 500kb of the SNP were studied: TMEM18, FAM150B, SH3YL1, ACP1, and SNTG2.

Therefore in total, 57 human genes were chosen for study.

4.5.2 Identification of *Drosophila* orthologues

The work in section 4.5.2 was done with Dr Loraine Tung. The *Drosophila* orthologues of each of the 57 human genes were identified using the ENSEMBL orthologue tool (Zerbino et al., 2018) and FlyBase BLAST searches of protein sequences (Altschul et al., 1990; Gramates et al., 2017). 45 out of the 57 human genes (79%) had at least 1 *Drosophila* orthologue, giving a total of 58 fly genes, as summarised in figure 4.2. Each GWAS SNP, its position, the nearest gene(s), and their *Drosophila* orthologues are all shown in table 4.1.

4.5.3 UAS_G-siRNA lines to target *Drosophila* genes

A UAS_G-siRNA line was available for each of the 58 *Drosophila* genes, whereas LoF lines were only available for 39 of the genes (table 2.1). This reinforced the decision made in chapter 3 to use RNAi as the primary tool for screening genes in *Drosophila*.

Of the siRNA lines ordered, two (*boss* and *sbb* which are the orthologues of GPRC5B and ZNF608 respectively) died in quarantine - these lines were re-ordered but suffered the same fate. In addition there was a problem with ordering the siRNA line for the SH2B1 orthologue *Lnk*, but fortunately there are already several papers which examine energy homeostasis phenotypes related to perturbation of this fly gene (for example Slack et al. (2010) and Song et al. (2010)).

Three of the GD lines used in this chapter (V14, VV22, and VV46) contain the UAS_G-siRNA insert on the X chromosome. As explained in section 3.6.2, these lines were delivered from VDRC as a compound X stock, and so were crossed onto an FM7 balancer prior to phenotyping.

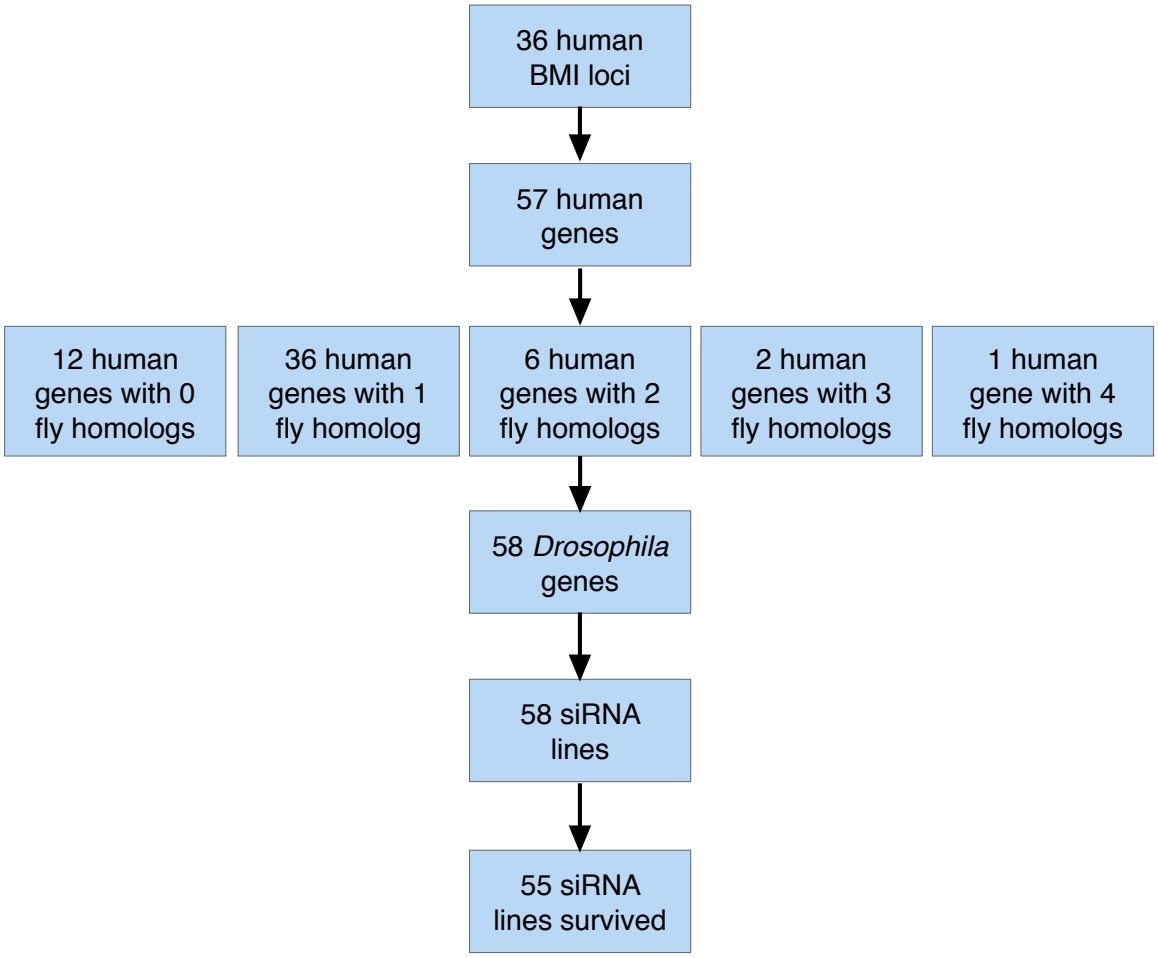


Fig. 4.2 79% of the human genes from GWAS BMI loci had at least one orthologue in *Drosophila*, resulting in acquisition of UAS_G-siRNA lines to study 55 *Drosophila* genes

Table 4.1 Genes within loci associated with human BMI and their *Drosophila* orthologues

SNP name	Chr	Position (bp)	Nearest gene	Nearby genes	*Drosophila* orthologue(s)	Homology (%)
rs1555543	1	96,717,385	PTBP2	-	*heph*	52
rs543874	1	176,156,103	SEC16B	-	*dSec16*	17
rs887912	2	59,156,381	FANCL	-	*Fancl*	22
rs2890652	2	142,676,401	LRP1B	-	*Lrp1*	36
rs13078807	3	85,966,840	CADM2	-	*sns*	29
rs9816226	3	187,317,193	ETV5	-	*Ets96B*	31
rs13107325	4	103,407,732	SLC39A8	-	*Zip71B*	29
rs261967	5	95,850,250	PCSK1	-	*fur1, fur2, rab2*	50,48, 48
rs4836133	5	124,360,002	ZNF608	-	*sbb*	17
rs9356744	6	20,685,255	CDKAL1	-	*CG6550*	59
rs206936	6	34,410,847	NUDT3	-	*Aps*	53
rs987237	6	50,911,009	TFAP2B	-	*TfAP-2*	41
rs10968576	9	28,404,339	LRRN6C	-	*trn, caps, fili, CG14762*	21, 23, 24, 29
rs11142387	9	72,998,332	KLF9	-	*sp1, cabot*	64, 31
rs4929949	11	8,561,169	RPL27A	-	*RpL27A*	66
rs10767664	11	27,682,562	BDNF	-	*NT1*	24
rs3817334	11	47,607,569	MTCH2	-	*CG10920, Mtch*	31, 35
rs7138803	12	48,533,735	FAIM2	-	*Nmda1*	42
rs4771122	13	26,918,180	MTIF3	-	*mIF3*	32
rs11847697	14	29,584,863	PRKD1	-	*PKD*	52
rs10150332	14	79,006,717	NRXN3	-	*Nrx-1*	36
rs12444979	16	19,841,101	GPRC5B	-	*boss*	28
rs29941	19	39,001,372	KCTD15	-	*twz*	46
rs1558902	16	52,361,075	FTO	-	-	-
rs571312	18	55,990,749	MC4R	-	-	-
rs1514175	1	74,764,232	TNNI3K	-	-	-
rs3810291	19	52,260,843	TMEM160	-	-	-
rs12597579	16	20246545	GP2	-	-	-
rs2867125	2	612,827	TMEM18	-	*Tmem18*	40
				SH3YL1	*RtGEF*	50
				SNTG2	*Syn2*	39
				ACP1	*primo-1*	46
				FAM150B	-	-
rs2815752	1	72,585,028	NEGR1	-	*Lac, DIP-iota, Ama*	24, 26, 24
				ZRANB2	*CG3732*	45
				NKX6.1	*HGTX*	39
				PTGER3	-	-
				LRRIQ3	-	-
rs7359397	16	28,793,160	SH2B1	-	*Lnk*	26
				APOB48R	-	-
				SULT1A2	*St2*	29
				AC138894.2	-	-
				ATXN2L	*Atx2*	8
				SBK1	*CG4945, PKN*	27, 25
				TUFM	*mEFTu1*	60
rs713586	2	25,011,512	RBJ	-	*rab21*	32
				ADCY3	*Ac3*	39
				POMC	-	-
rs2241423	15	65,873,892	MAP2K5	-	*dsor1*	42
				LBXCOR1	*fuss*	29
rs2287019	19	50,894,012	QPCTL	-	*QC, isoQC*	40, 47
				GIPR	*Dh44-R1, Dh44-R2*	30, 30
rs2112347	5	75,050,998	FLJ35779	-	-	-
				HMGCR	*Hmgcr*	47
				COL4A3BP	*cert*	44
rs10938397	4	44,877,284	GNPDA2	-	*Oscillin*	74
				GABRG1	*CG8916, Grd*	32, 37

4.6 Expression of BMI GWAS genes in the *Drosophila* CNS

The CNS has a key role in the regulation of energy homeostasis and feeding behaviour (see chapter 1). Published gene expression microarray data has shown that expression of genes at the GWAS BMI loci is enriched in the CNS (Locke et al., 2015), in contrast to GWASs of waist-to-hip ratio where the genes show enriched expression in adipose tissue (Shungin et al., 2015).

Therefore, the expression pattern of the GWAS BMI *Drosophila* homologues was examined using published microarray data which shows RNA levels in different *Drosophila* tissues and developmental stages (Graveley et al., 2011). According to this data, of the 58 *Drosophila* genes selected for study 55 are expressed in the larval, pupal and/or adult CNS. Of the other 3 genes, *Grd* was not detected anywhere in the body, and *CG10920* and *Ets96B* were only expressed in the *Drosophila* testes.

4.7 Lethality of BMI GWAS genes in *Drosophila*

All of the *Drosophila* lines with neuron-specific RNAi of the GWAS BMI genes were viable as adults. However, the same is not true for the whole-body perturbations, and the lethality/viability of the each gene is shown in figure 4.3.

For the 35 RNAi lines with a GD background, 18 (51%) did not produce viable male adults with ubiquitous RNAi. Some of these lethal crosses were bred at 18^{o}C because GAL4 is less active at this temperature (Duffy, 2002), but the crosses were still not viable. For the siRNA lines with a KK background, 10 out of 19 genes did not produce viable male adults, although interpretation of the KK results is complicated by the fact that over-expression of *tiptop* in some lines also causes lethality (chapter 3.6.2). 4 genes showed male-specific lethality: *fur2* (PCSK1), *PKD* (PRKD1), *Lac* and *DIP-iota* (both NEGR1).

Of the 39 BMI GWAS *Drosophila* genes available as a LoF line, 23 (59%) were not viable as homozygous adults (these stocks were maintained on balancer chromosomes).

In total, the viability of 38 genes was tested with both the RNAi and the LoF models. Of these, 13 were lethal with both models, and 5 were viable with both. In addition, 9 genes were viable with RNAi but not LoF, and 7 GD lines (plus 4 KK lines) were viable with LoF but not RNAi.

Same in both models						
***Drosophila* gene**	**Human gene**	**RNAi line**			**LoF line**	
heph	PTBP2	VV15	GD	lethal	B1	lethal
dsor1	MAP2K5	VV18	GD	lethal	B3	lethal
RpL27A	RPL27A	VV23	GD	lethal	B4	lethal
trn	LRRN6C	VV7	GD	lethal	B2	lethal
HGTX	NKX6-1	VV8	GD	lethal	B14	lethal
fur1	PCSK1	VV9	GD	lethal	B33	lethal
Nmda1	FAIM2	VV28	KK	lethal	B32	lethal
Hmgcr	HMGCR	VV29	KK	lethal	B8	lethal
Atx2	ATXN2L	VV30	KK	lethal	B16	lethal
PKD	PRKD1	VV58	KK	male lethal	B30	lethal
caps	LRRN6C	VV10	GD	semi-lethal	B9	lethal
rab2	PCSK1	VV16	GD	semi-lethal	B15	lethal
RtGEF	SH3YL1	VV31	GD	semi-lethal	B34	lethal
Dh44-R2	GIPR	VV42	GD	viable	B31	viable
Grd	GABRG1	VV55	GD	viable	B24	viable
Mtch	MTCH2	VV43	GD	viable	B22	viable
Zip71B	SLC39A8	VV51	GD	viable	K6	viable
Aps	NUDT3	VV48	KK	viable	K1	viable

Discordant results						
***Drosophila* gene**	**Human gene**	**RNAi line**			**LoF line**	
Nrx-1	NRXN3	VV47	GD	lethal	B17	viable
QC	QPCTL	VV50	GD	lethal	B23	viable
CG14762	LRRN6C	VV56	GD	lethal	B29	viable
Oscillin	GNPDA2	VV27	KK	lethal	H5	viable
Syn2	SNTG2	VV52	KK	lethal	B27	viable
Dh44-R1	GIPR	VV41	KK	lethal	B20	viable
sns	*CADM2*	VV40	KK	lethal	B25	viable
Ama	NEGR1	VV46	GD	semi-lethal	K2	viable
Lac	NEGR1	VV44	GD	male-lethal	B12	viable
fur2	PCSK1	VV49	GD	male-lethal	K3	viable
DIP-iota	NEGR1	VV45	GD	male-lethal	B19	viable
Fancl	FANCL	VV13	GD	viable	K4	lethal
tmem18	TMEM18	VV21	GD	viable	K5	lethal
Fili	LRRN6C	VV22	GD	viable	B28	lethal
mEFTu1	TUFM	VV24	GD	viable	B10	lethal
cert	COL4A3BP	VV11	GD	viable	H4	lethal
Ac3	ADCY3	VV14	GD	viable	B35	lethal
Lrp1	LRP1B	VV17	GD	viable	B13	lethal
isoQC	QPCTL	VV25	KK	viable	H1	lethal
CG3732	ZRANB2	VV26	KK	viable	H2	lethal

Only tested in one model						
***Drosophila* gene**	**Human gene**	**RNAi line**			**LoF line**	
Cabot	KLF9	V10	GD	viable	-	not tested
Ets96B	ETV5	V14	GD	viable	-	not tested
Sec16	SEC16B	V15	GD	viable	-	not tested
CG6550	CDKAL1	V8	GD	viable	-	not tested
Sp1	KLF9	V9	GD	viable	-	not tested
CG4945	SBK1	VV59	GD	viable	-	not tested
Rab21	RBJ	V11	KK	viable	-	not tested
NT1	BDNF	V16	KK	viable	-	not tested
fuss	LBXCOR1	V3	KK	viable	-	not tested
CG8916	GABRG1	V7	KK	viable	-	not tested
mIF3	MTIF3	VV60	KK	viable	-	not tested
St2	SULT1A2	V4	GD	lethal	-	not tested
TfAP2	TFAP2B	VV19	GD	lethal	-	not tested
twz	KCTD15	VV54	KK	lethal	-	not tested
PKN	SBK1	V5	KK	lethal	-	not tested
CG10920	MTCH2	V1	KK	lethal	-	not tested
sbb	ZNF608	VV57	GD	not tested	B11	lethal
boss	GPRC5B	V13	GD	not tested	-	not tested

Fig. 4.3 The viability/lethality of flies with either whole body RNAi of, or LoF of, genes homologous to human GWAS BMI genes. Lethal KK lines are highlighted in yellow because this phenotype may be caused by *tiptop* over-expression rather than by the gene of interest.

4.8 Energy homeostasis phenotypes of BMI GWAS genes

Lines screened

In section 4.5, 58 *Drosophila* homologues were identified of 44 human genes which are located in genetic loci suggested by GWAS to influence BMI. The *Drosophila*-based screen that was validated in chapter 3 was used to phenotype *Drosophila* adults with RNAi knockdown of these BMI GWAS genes. act-GAL4 was used to direct constitutive whole-body RNAi knockdown of BMI GWAS genes — 27 of these crosses produced viable male adults which were assayed. elav-GAL4 was used to direct neuron-specific RNAi knockdown of BMI GWAS genes — all of these crosses were viable. As auxiliary studies, larval feeding motivation was assessed in the 39 LoF lines, and adult feeding motivation was measured in the top RNAi hits.

Defining the top hit genes

For each individual assay, a hit was defined as a gene in which knockdown resulted in a statistically significant phenotype ($p<0.05$). For the overall score, a hit was defined as >0.7, and a strong hit as >0.8, based on the results of the positive and negative control genes in chapter 3.

Leanness and decreased food intake phenotypes may truly result from the manipulation of genes important in energy balance, but may also be linked to sickness of the animal due to the genetic manipulation (Yazdi et al., 2015). These phenotypes (referred to in this thesis as "decrease phenotypes") were therefore deemed less important than "increase phenotypes" (i.e. fatness, increased food intake), even if they showed a greater overall score.

Lines affected by *tiptop* over-expression

As was shown in chapter 3.6.2, some KK RNAi lines are affected by over-expression of the *tiptop* gene and it is important to control for this in order to avoid false positive results.

Nine of the siRNA lines with a KK background (CG10920, PKN, Oscillin, Hmgcr, atx2, sns, Dh44-R1, Syn2, twz) were lethal with whole body RNAi, marking them as potentially affected by *tiptop*. Therefore, all of these lines were compared to both KK-elavGAL4 and KK_{tiptop}-elavGAL4 as a background. Since the purpose of this screen is candidate gene selection, it is better to have false negative results than false positives. Therefore, the smaller phenotype was chosen as the background for each line, as summarised in figure 4.4.

Three siRNA lines (CG10920, PKN, Syn2) were assigned to a KK_{tiptop} background because this gave the smaller phenotype. Two of these were not tested as LoF lines, but a *Syn2* LoF line was viable which agrees with the theory that the lethality observed is caused by *tiptop*.

The other six siRNA lines (Oscillin, Hmgcr, atx2, sns, Dh44-R1, twz) were assigned to a KK background because this gave the smaller phenotype. *Hmgcr, atx2*, and *sns* LoF lines were not viable, which agrees with the idea that the lethality seen with RNAi is caused by the gene itself, and not by *tiptop*. *twz* was not tested as a LoF line. The other two genes (*Oscillin* and *Dh44-R1*) are viable as LoF lines, but their phenotype is very strongly decreased compared to KK_{tiptop}, so even if this background was chosen, the genes would be considered unhealthy and not worthy of further study.

Line	Background	Fly Gene	Human Gene	CAFE	Dye Food	Mass	Glucose	Climbing	Starvation	Score	LoF line
VV27	KK	*Oscillin*	GNPDA2	0.7122	0.8363	0.0534	0.1759	0.3341	0.6564	0.50	viable
	tiptop			**0.0125**	0.1023	**0.0007**	0.1682	0.0589	0.4421	0.92	
VV41	KK	*Dh44-R1*	GIPR	0.1591	0.6990	0.5933	0.0913	**0.0192**	0.1658	0.65	viable
	tiptop			0.3037	**0.0366**	**0.0115**	0.0112	0.9124	0.0990	0.81	
VV29	KK	*Hmgcr*	HMGCR	0.7354	0.2351	**0.0177**	0.3297	**0.0360**	0.3538	0.66	lethal
	tiptop			0.1278	**0.0077**	**0.0001**	**0.0027**	0.2589	0.7092	0.90	
VV30	KK	*Atx2*	ATXN2L	0.6370	0.1066	0.9199	**0.0104**	0.6847	0.9653	0.50	lethal
	tiptop			**0.0155**	**0.0071**	**0.0163**	0.0655	**0.0201**	0.6663	0.95	
VV40	KK	*sns*	CADM2	0.4087	0.3620	0.5579	0.1771	0.1011	0.3645	0.64	lethal
	tiptop			0.2226	**0.0113**	0.0599	**0.0054**	0.5767	0.2245	0.86	
VV54	KK	*twz*	KCTD15	0.1705	0.3336	0.4272	0.2527	0.9324	0.4445	0.64	not tested
	tiptop			0.1783	**0.0202**	**0.0077**	**0.0047**	**0.0376**	0.8444	0.90	
V1	KK	*CG10920*	MTCH2	0.8032	**0.0433**	**0.0094**	**0.0311**	0.1302	0.6505	0.71	not tested
	tiptop			**0.0245**	1.0000	0.1605	0.5207	0.0658	0.9329	0.60	
V5	KK	*PKN*	SBK1	0.2889	0.0560	**0.0318**	**0.0058**	0.5654	0.2987	0.83	not tested
	tiptop			**0.0034**	0.8421	0.4522	0.6849	**0.0079**	0.6792	0.58	
VV52	KK	*Syn2*	SNTG2	0.1213	0.1770	**0.0341**	0.3209	0.8005	0.8430	0.75	viable
	tiptop			0.3600	0.4219	0.5681	**0.0260**	0.1470	0.7725	0.63	

Fig. 4.4 KK lines which were lethal with whole body RNAi were compared to both KK-elavGAL4 and KK_{tiptop}-elavGAL4 as potential backgrounds, and the smaller phenotype chosen (orange). Red is an increase compared to control, blue is a decrease and significant values are highlighted.

Results

A summary of the phenotypes of the GWAS BMI lines in each assay, as well as the overall scores, are shown in figure 4.5 for neuronal RNAi, and figure 4.6 for whole body RNAi. The full phenotypes are shown in figures 4.7 to 4.12. In each graph, the genes which show the strongest overall phenotype compared to background (a score >0.8) are highlighted in purple. Conversely, the genes which are the most similar to background (a score <0.45) are highlighted in yellow. The auxiliary feeding motivation studies are in figures 4.14 and 4.15.

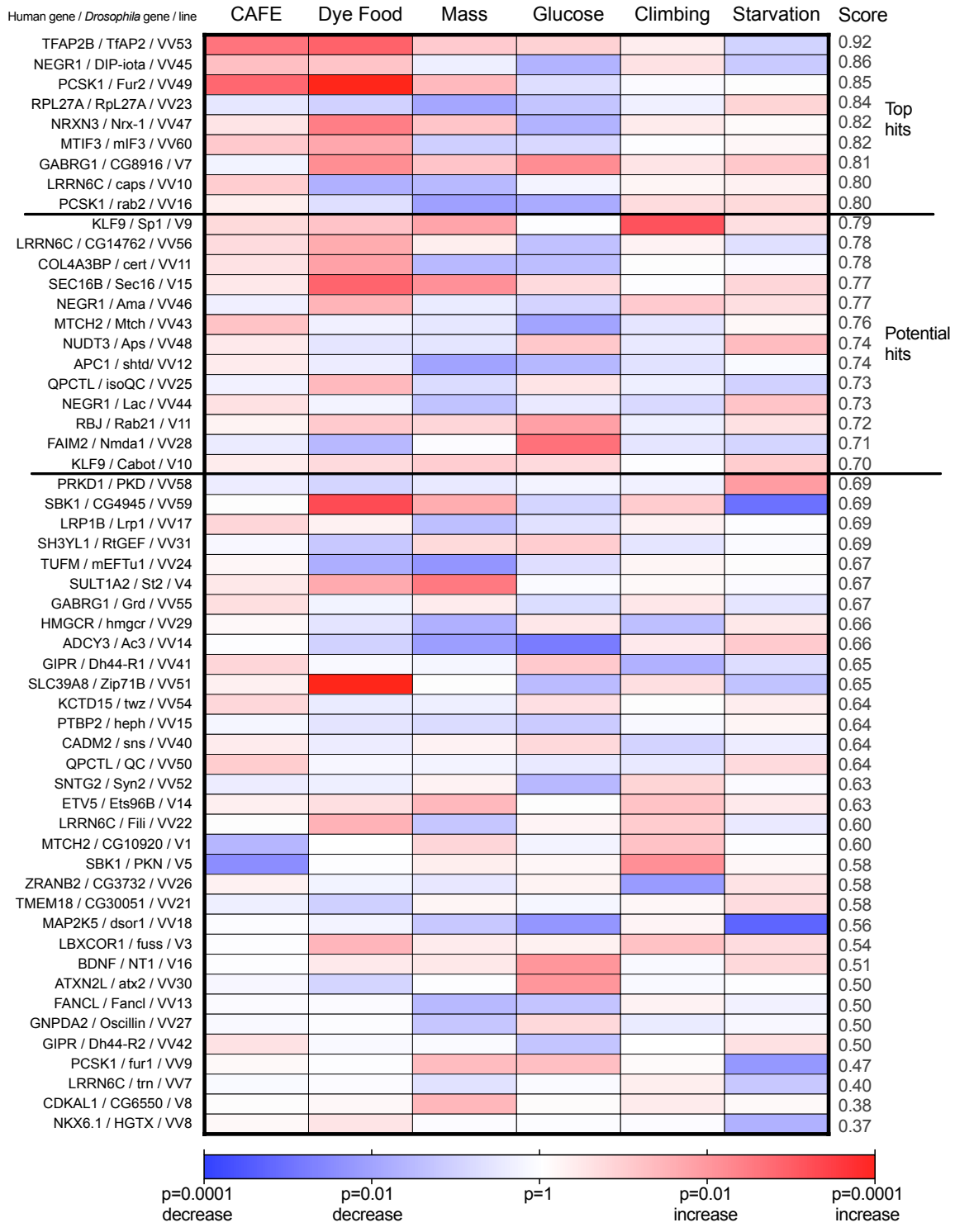


Fig. 4.5 Summary of the phenotype and overall score of flies with neuronal RNAi knockdown of GWAS BMI gene homologues

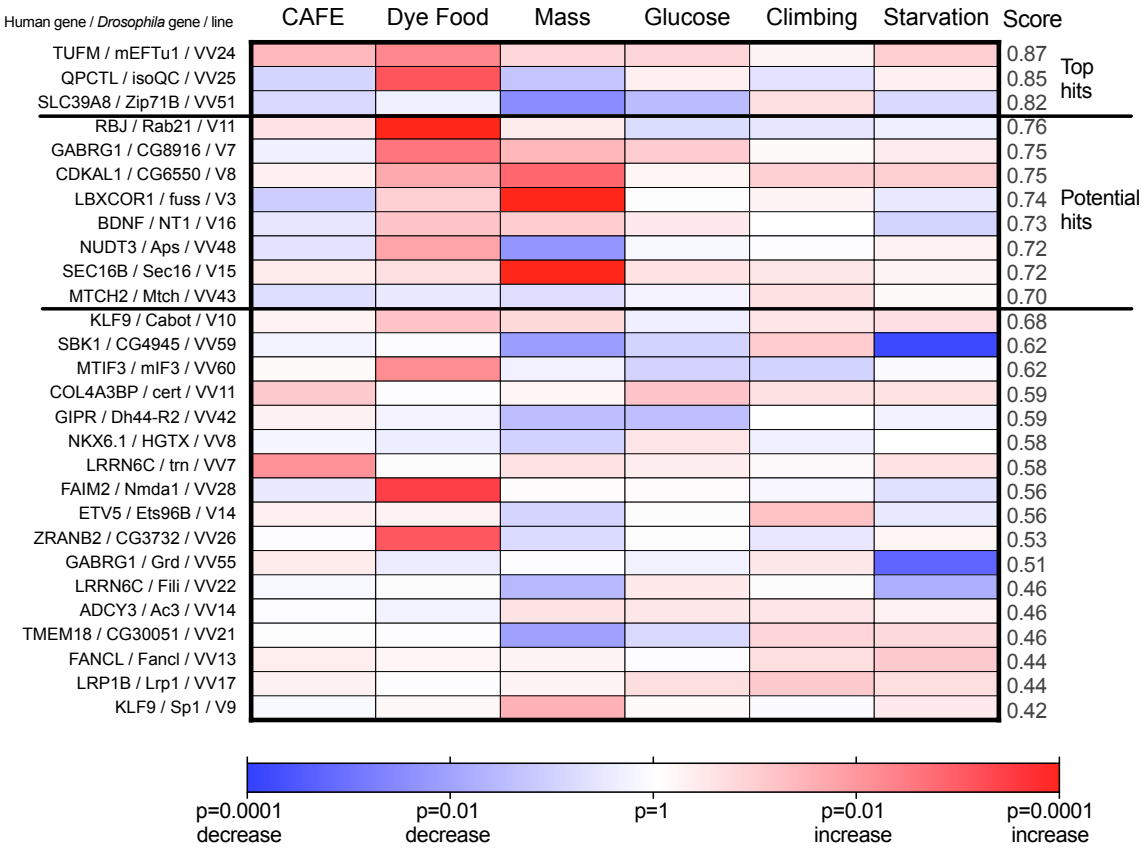


Fig. 4.6 Summary of the phenotype and overall score of flies with whole body RNAi knockdown of GWAS BMI gene homologues

CAFE assay

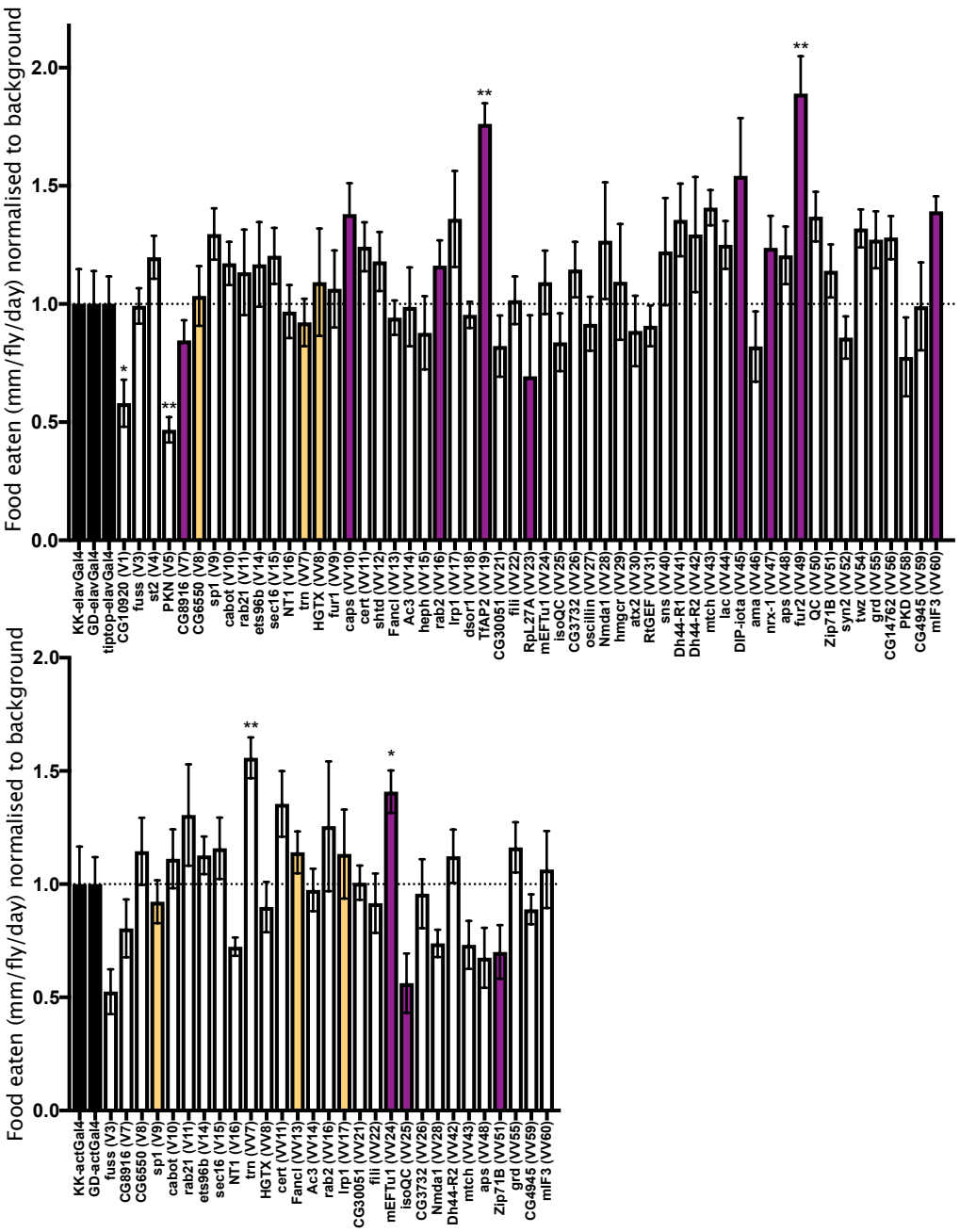


Fig. 4.7 CAFE food intake by flies with (top) neuron-specific and (bottom) whole-body RNAi knockdown of GWAS BMI gene homologues. Background lines are black, the overall top-scoring genes are purple, and bottom-scoring genes are yellow. Mean ± is plotted of 5 repeats each containing 8 flies. Results were compared to appropriate background by Student's t-test. * $p<0.05$, ** $p<0.01$

Over-feeding dye assay

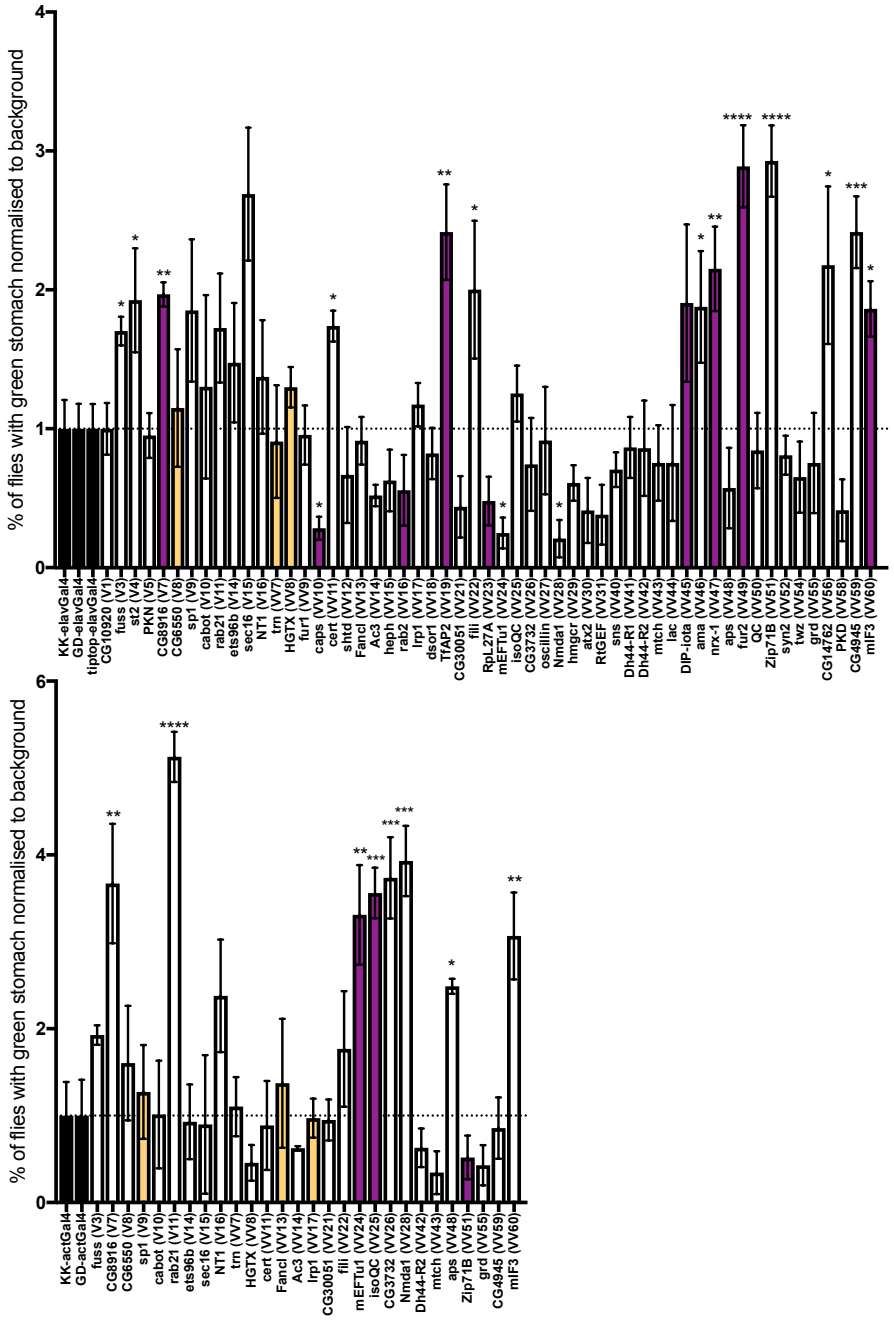


Fig. 4.8 Dye food over-feeding by flies with (top) neuron-specific and (bottom) whole-body RNAi knockdown of GWAS BMI gene homologues. Background lines are black, the overall top-scoring genes are purple, and bottom-scoring genes are yellow. Mean ± is plotted of 5 repeats each containing 15 flies. Results were compared to appropriate background by Student's t-test. * p<0.05, ** p<0.01, *** p<0.001, **** p<0.0001

Wet mass assay

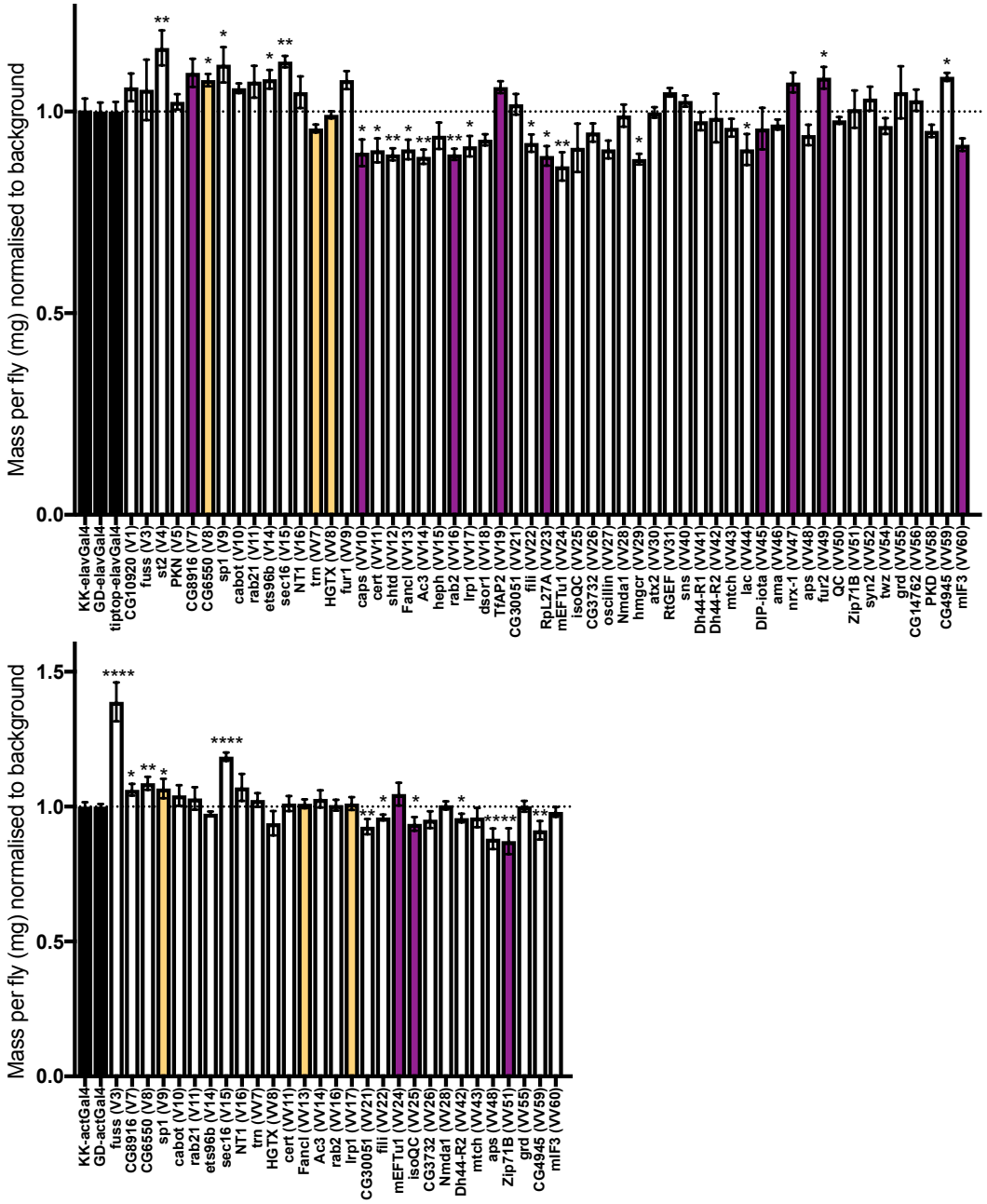


Fig. 4.9 Wet mass of flies with (top) neuron-specific and (bottom) whole-body RNAi knockdown of GWAS BMI gene homologues. Background lines are black, the overall top-scoring genes are purple, and bottom-scoring genes are yellow. Mean ± is plotted of 5 repeats each containing 15 flies. Results were compared to appropriate background by Student's t-test. * p<0.05, ** p<0.01, *** p<0.001, **** p<0.0001

Glucose assay

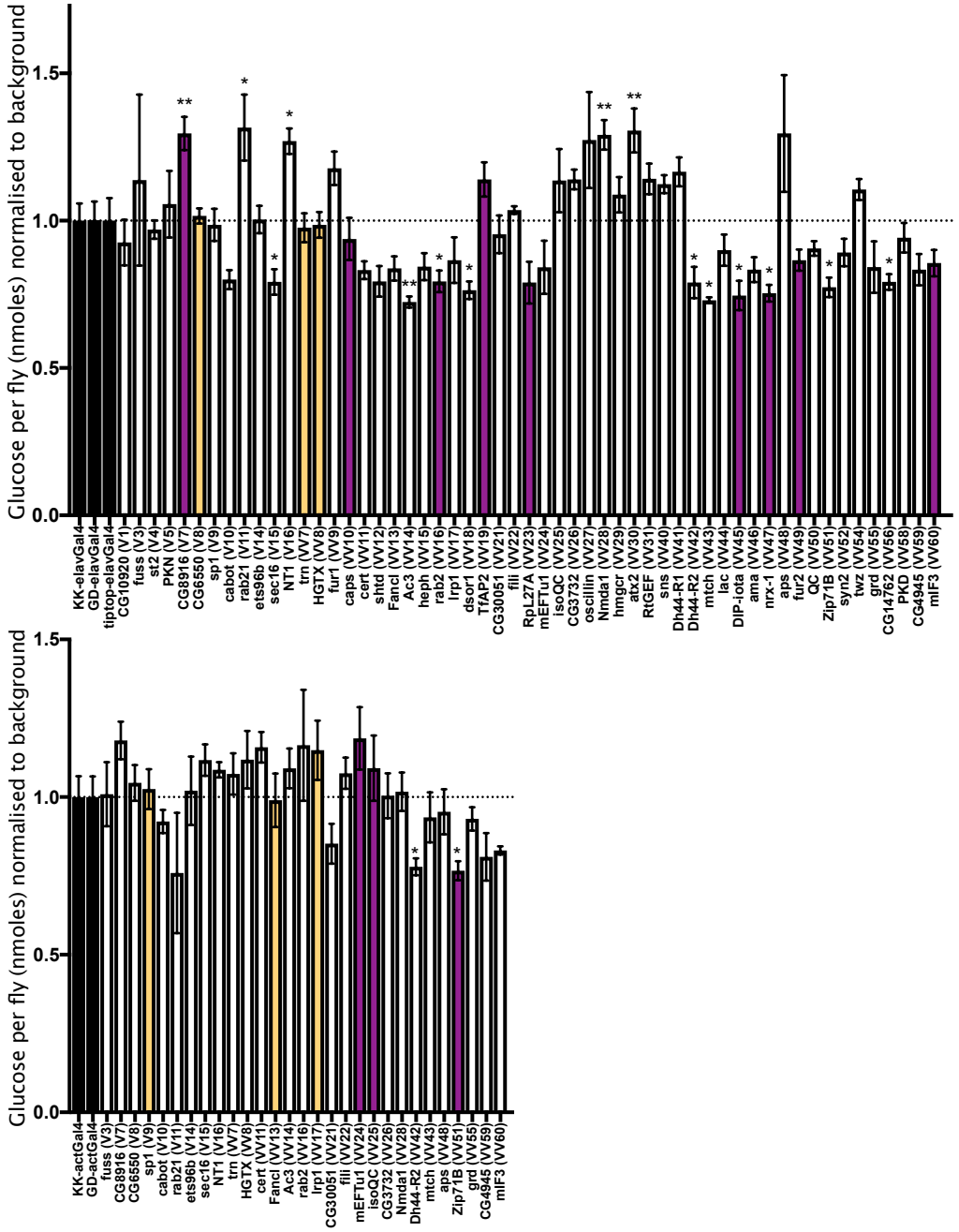


Fig. 4.10 Glucose levels in flies with (top) neuron-specific and (bottom) whole-body RNAi knockdown of GWAS BMI gene homologues. Background lines are black, the overall top-scoring genes are purple, and bottom-scoring genes are yellow. Mean ± is plotted of 5 repeats each containing 15 flies. Results were compared to appropriate background by Student's t-test. * p<0.05, ** p<0.01

Climbing assay

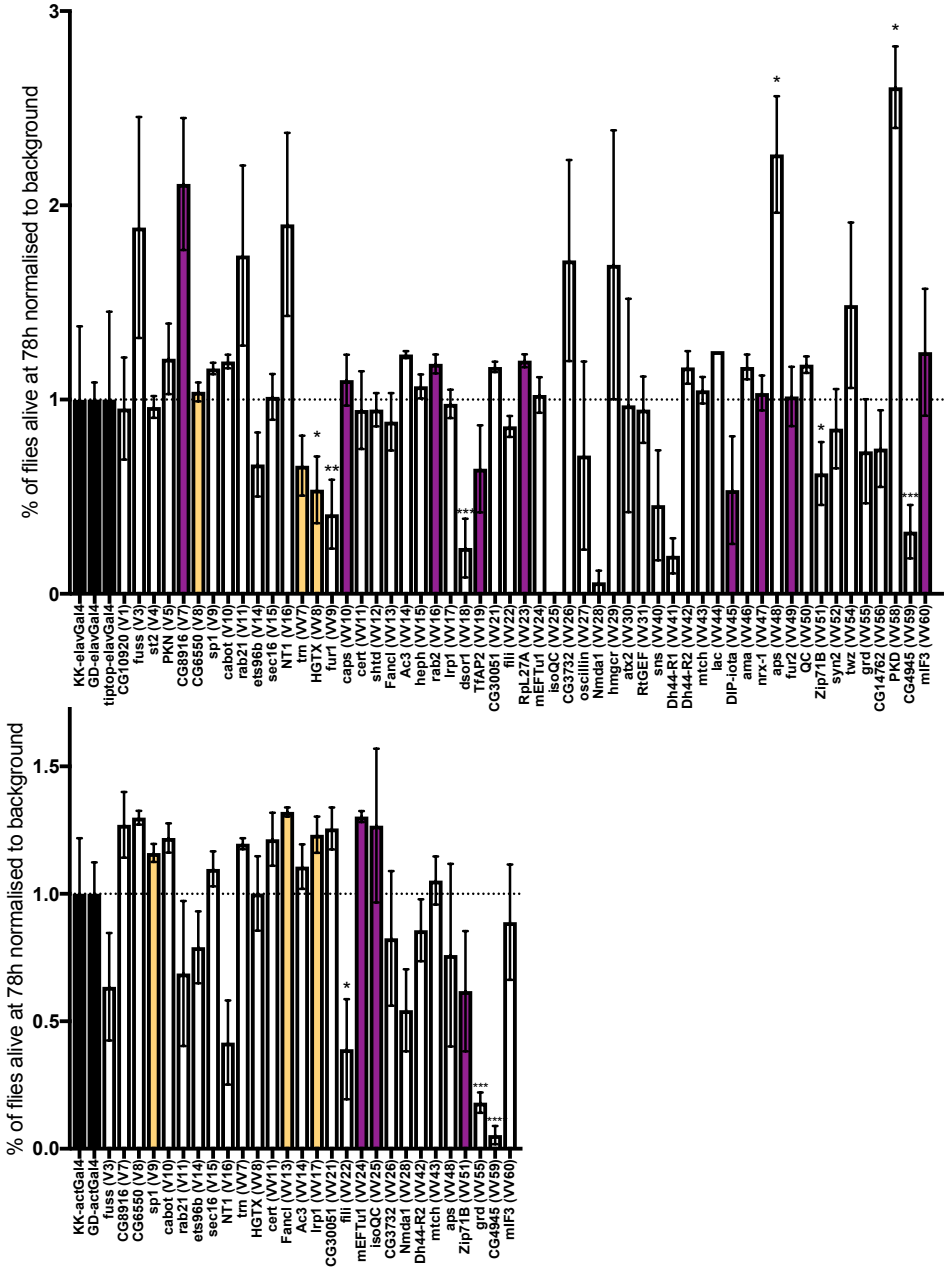


Fig. 4.11 Climbing ability of flies with (top) neuron-specific and (bottom) whole-body RNAi knockdown of GWAS BMI gene homologues. Background lines are black, the overall top-scoring genes are purple, and bottom-scoring genes are yellow. Mean ± is plotted of 5 repeats each containing 15 flies. Results were compared to appropriate background by Student's t-test. * p<0.05, ** p<0.01, *** p<0.001

Starvation resistance assay

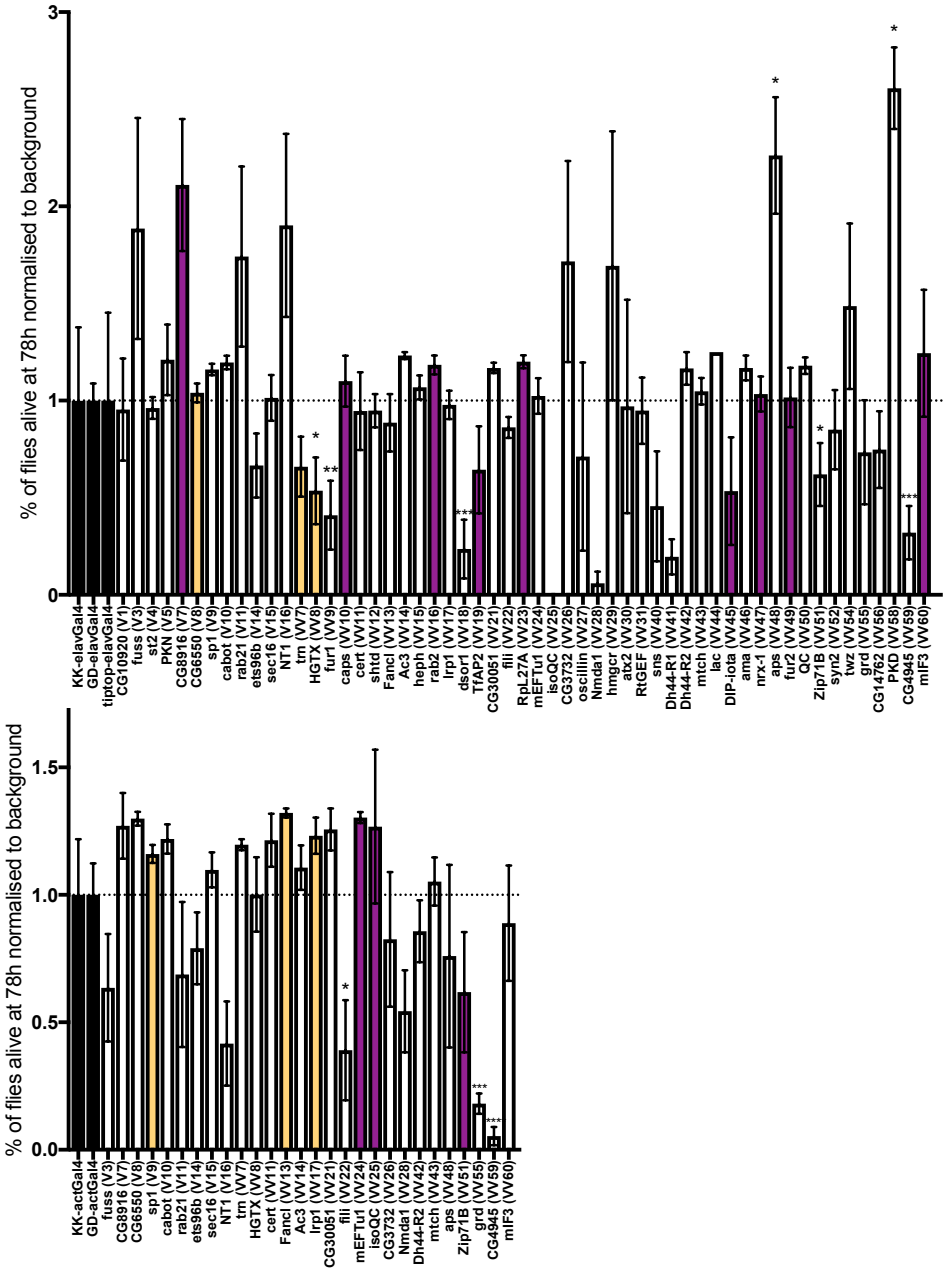


Fig. 4.12 Survival during starvation of flies with (top) neuron-specific and (bottom) whole-body RNAi knockdown of GWAS BMI gene homologues. Background lines are black, the overall top-scoring genes are purple, and bottom-scoring genes are yellow. Mean ± is plotted of 5 repeats each containing 15 flies. Results were compared to appropriate background by Student's t-test. * p<0.05, ** p<0.01, *** p<0.001, **** p>0.0001

Auxiliary study: larval feeding motivation

The LoF lines which were viable as homozygous adults were assayed by Dr Loraine Tung.

The LoF lines which were lethal as adults may be viable as larvae, but studying these larvae required a slightly different approach. These stocks were maintained on balancer chromosomes which have easily visible markers in adult flies, but not in larvae. Therefore, the balancers were replaced (by crossing) with YFP-tagged FM7, GFP-tagged CyO, or GFP-tagged TM3 as appropriate. The GFP was expressed throughout the larval body whereas the YFP was expressed in the larval mouth hook. This allowed the homozygous larvae to be separated from the balanced heterozygous larvae by their absence of fluorescence (see figure 4.13). Of the 23 lethal LoF lines, this YFP/GFP-based sorting showed that only 8 were viable as third instar larvae. Homozygotes from these 8 lines were selected and used to study larval feeding motivation.

The results are shown in figure 4.14. *Nrx-1* and *Grd* show significant increases in larval feeding motivation. Also, most genes show a significant difference in behaviour between the fed and the fasted state, but a few (*Nrx-1, DIPiota, Dh44-R2, fur2 RtGEF, Ac3*) do not.

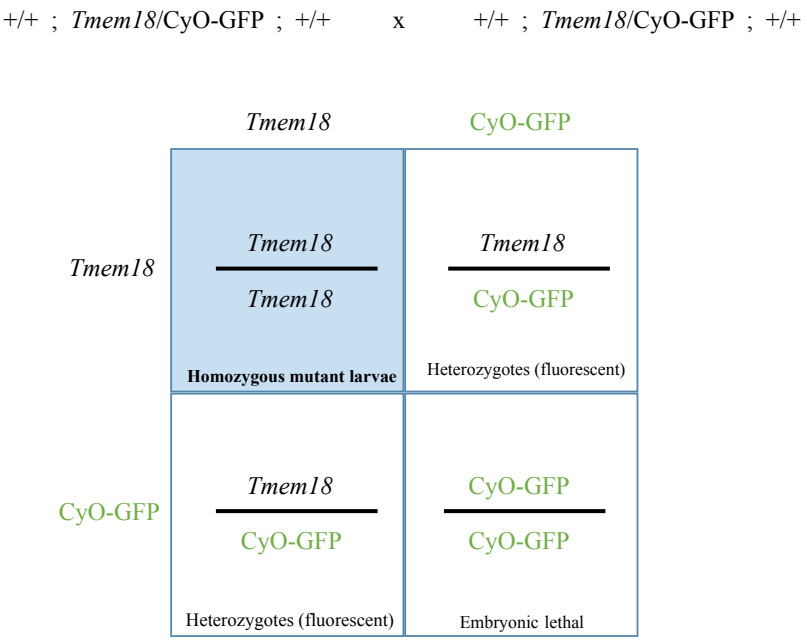


Fig. 4.13 Homozygous mutant larvae were selected for assay by their absence of fluorescence.

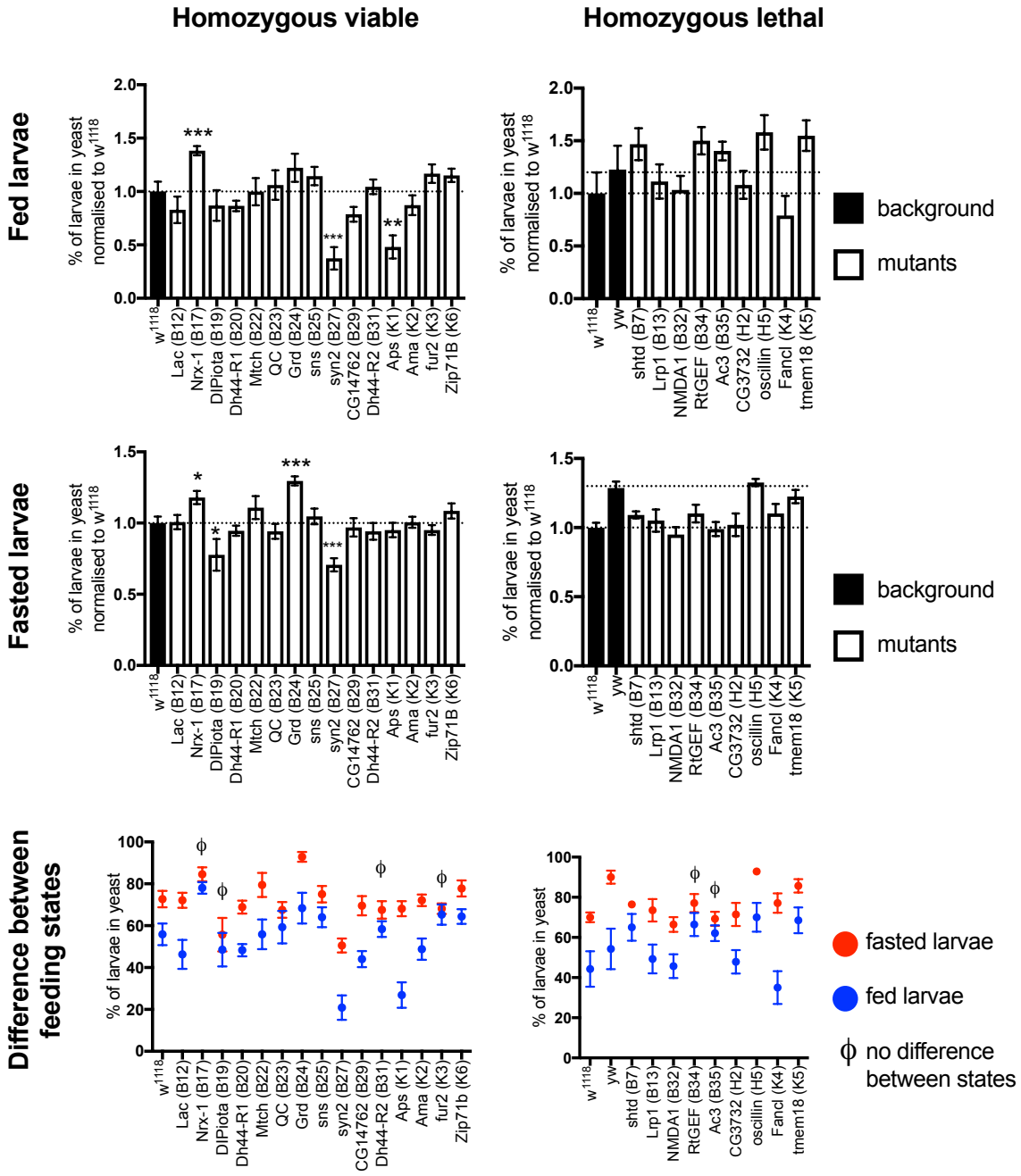


Fig. 4.14 Feeding motivation of *Drosophila* larvae with LoF of BMI GWAS genes in the fed and fasted states. Mean ± SEM is plotted. All statistical comparisons were Student's t-test. Homozygous viable lines were assayed by Dr Loraine Tung. For fed (top) and fasted (middle) larvae, * p<0.05, ** p<0.01, *** p<0.001 compared to background. For the difference graphs (bottom), ϕ indicates that there is no significant difference between fed and fasted larvae for the gene.

Auxiliary study: adult feeding motivation

In chapter 3, an assay was developed to test a fly's motivation to approach food despite the presence of a repellent. The *Drosophila* genes with the top 5 scores from the screen (*TfAP-2, DIPiota, fur2, Nrx1 mEFTu1*) were tested with this assay. The gene with the lowest score from the screen (*HGTX*) was also included as a negative control. The results are shown in figure 4.15. A significant change was only detected for *TfAP-2* but this was not driven by differences in the food/repellent zone and so the relevance of this result is unclear.

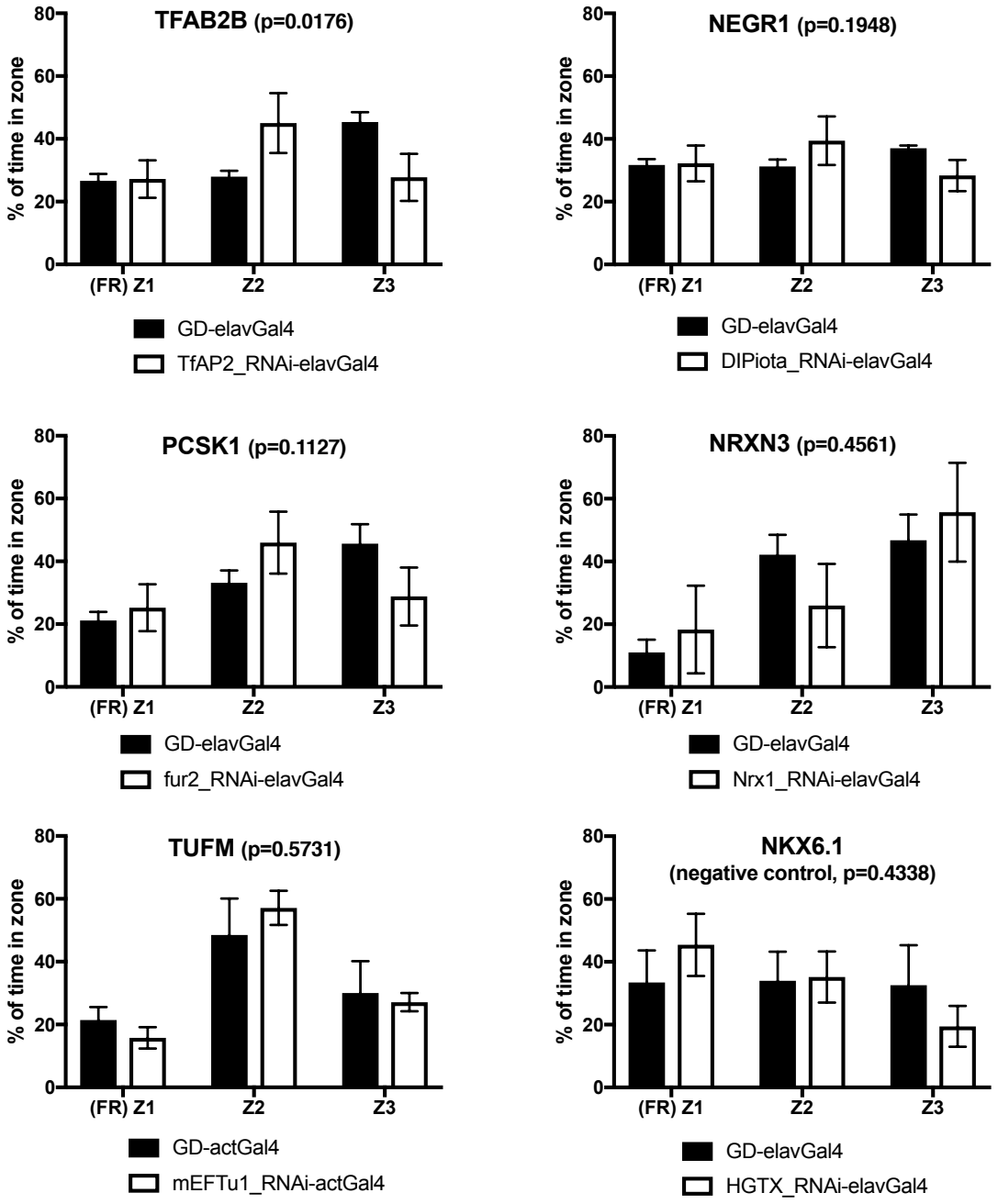


Fig. 4.15 Time spent by *Drosophila* close to yeast despite the presence of benzaldehyde repellent . Mean ± SEM is plotted of 7 repeats. Data was analysed by two-way ANOVA. FR = food + repellent.

4.9 Comparison of *Drosophila* screen results to published literature

As has been mentioned previously, Baranski et al. (2018) carried out a screen of BMI GWAS genes looking for changes in levels of TAG. Despite using a different initial list of human genes, and a different computational tool to select their *Drosophila* orthologues, 17 *Drosophila* genes were screened in common.

Their preference for use of KK lines and ours for use of GD lines means that for 12 of these genes, a different siRNA line was used. Therefore their results provide a method to independently verify the results obtained here for these 17 genes (although Baranski et al. used RNAi in both the neurons and the fat body (cg-GAL4) compared to just the neurons (elav-GAL4) here).

As can be seen in table 4.2, Baranski et al. observed increased TAG for 5 of the 17 genes (*CG4945, Nrx-1, Sec16, Aps* and *cert*). Here, a score >0.7 was seen for 4 out of these genes (all except *CG4945*). Further, for three of the genes (*lac, isoQC* and *mtch*), Baranski et al. did not see any TAG phenotype, but because multiple assays were used here, the screen was able to detect more subtle phenotypes, suggesting increased sensitivity. Finally, they found that 2 of the genes (*RpL27A and EfTuM*) gave a lethal phenotype (and overall 13% of their genes were lethal) whereas all of our crosses were viable.

Table 4.2 Results obtained in this thesis generally agree with those published by Baranski et al. (2018). Scores >0.7 are highlighted in red. *The same siRNA line was used

Human gene	Fly gene	Adult TAG (Baranski et al., 2018)	Neuronal RNAi score
SBK1	*CG4945* *	fatter	0.69
NRXN3	*Nrx-1* *	fatter	0.82
SEC16B	*Sec16*	fatter	0.77
NUDT3	*Aps* *	fatter	0.74
COL4A3BP	*cert*	fatter	0.78
SLC39A8	*CG10006*	wt	0.65
SKOR1	*fuss* *	wt	0.54
NEGR1	*Lac*	wt	0.73
TMEM18	*Tmem18*	wt	0.58
LRP1B	*LRP1*	wt	0.69
ATXN2L	*Atx2* *	wt	0.50
QPCTL	*isoQC*	wt	0.73
MTCH2	*Mtch*	wt	0.76
GPRC5B	*boss*	wt	parental line died
KTCD15	*twz*	wt	0.64
RPL27A	*RpL27A*	lethal	0.84
TUFM	*EfTuM*	lethal	0.67

4.10 Phenotype of genes not expressed in the CNS

As discussed in section 4.6, *Grd, CG10920* and *Ets96B* are not expressed in the *Drosophila* CNS according to published microarray data. Therefore these genes can be used as negative controls for the neuronal RNAi. As would be expected, none of these three genes were a hit with neuronal RNAi: their scores were <0.7 and none of the individual assays were significant.

4.11 Phenotype of genes with altered backgrounds

For another three of the genes (*Ets96b*, *Fili Ama*), the siRNA transgene was on the X chromosome and so the stock was balanced (by VDRC) using a compound X. As explained in chapter 3, such lines cannot be used in the screen until the compound X is replaced with an FM7 balancer. However this process of changing the balancer may alter the background away from GD, making GD-GAL4 an inappropriate background to use for statistical analysis.

Of these three genes, *Ets96b* and *Fili* show no phenotype in the screen. The other gene, *Ama*, does have a strong phenotype, and so this result must be viewed with caution. *Ama* is an orthologue of NEGR1, and luckily this gene had several other orthologues which were also studied.

4.12 Interpretation of results according to human gene

The aim of this screen was to select genes which are top candidates for further study from an original list of human genes that have been associated with BMI by GWASs. The results of the *Drosophila*-based screen will now be presented in terms of the human genes tested. Where available, Baranski et al.'s results are included too (Baranski et al., 2018), because a TAG measurement was not included in the screen used here, although note that in some cases they used a different *Drosophila* orthologue of the same human gene.

4.12.1 Loci at which just one gene was tested

At 28 of the 36 SNP loci studied, only the closest gene was studied. Out of these single gene loci, 23 were testable in *Drosophila*. 14 genes showed a phenotype and 10 of these were considered worthy of further study.

Single gene loci that were not testable

Five loci could not be studied because no *Drosophila* orthologues were found: FTO, MC4R, TNN13K, TMEM160 and GP2. An additional two genes (ZNF608 and GPRC5B) could not be studied because the siRNA lines to target the *Drosophila* homologues *sbb* and *boss* died — they were re-ordered from VDRC but suffered the same fate.

Single gene loci that were not hits

At seven of the GWAS loci, the genes were not classed as hits in the *Drosophila*-based screen based on their score (<0.7). Further, none of these genes showed a significant result in more than one assay. These genes (and their *Drosophila* orthologues) were: ETV5 (*Ets96b*), FANCL (*Fancl*), CADM2 (*sns*), PRKD1 (*PKD*), PTBP2 (*heph*), LRP1B (*Lrp1*), and KCTD15 (*twz*). Baranski et al. did not find a TAG phenotype for their orthologues of ETV5, LRP1B or KCTD15, and did not test the other genes.

Four genes — RPL27A (*RpL27A*), FAIM2 (*Nmda1*), SLC39A8 (*Zip71B*) and MTCH2 (*Mtch*) — had scores which were indicative of a hit (<0.7). However, these scores were composed solely/mostly of decrease phenotypes which could be a specific energy homeostasis phenotype, but is indistinguishable from a general unhealthiness. Thus these four genes were not deemed worthy of further study.

Single gene loci that were potential hits

Four genes had scores between 0.7 and 0.8 composed of increase phenotypes, marking them as potentially worth investigation. BDNF (*NT1*) and CDKAL1 (*CG6550*) showed a phenotype only with ubiquitous RNAi, whereas NUDT3 (*Aps*) did so only with neuronal RNAi. In contrast, SEC16B (*dSec16*) is a hit with both RNAi paradigms. Baranski et al. found that their orthologues of both SEC16B and NUDT3 have increased TAG, and did not test the other two genes. Another study has shown that neuronal RNAi of *Aps* in *Drosophila* increased food intake but decreased starvation resistance and circulating glucose (Williams et al., 2015).

Single gene loci that were strong hits

KLF9 had scores between 0.7 and 0.8 for neuronal RNAi, and these were related to increase phenotypes, marking it as potentially worth investigation. However, because KLF9 has two *Drosophila* orthologues (*sp1* and *cabot*) which both show the same phenotype, this gene was considered a strong hit.

Four additional genes were highlighted as top hits based on their neuronal RNAi score being >0.8 and composed of increase phenotypes: TFAP2B (*TfAP-2*), NRXN3 (*Nrx-1*), PCSK1 (*fur2, rab2*) and LRRN6C (*caps, CG14672*). These were lethal with ubiquitous RNAi. PCSK1 and LRRN6C both have multiple homologs which are top scores, increasing confidence in these genes. In addition, *Nrx-1* showed a significant increase in larval feeding in both the fed and fasted states, and no difference in behaviour between these two states (i.e. fed larvae behaved as if they were fasted). Similarly, *fur2* larvae showed no difference in behaviour between the fed and fasted states. Baranski et al. found increased TAG for their NRXN3 orthologue, and increased food intake has been reported in flies with neuronal RNAi of *TfAP-2* (Williams et al., 2014).

MTIF3 (*mIF3*) was also considered to be a top hit, although it has a discordant phenotype with neuronal RNA: *Drosophila* show increased feeding but decreased measurements in the physiology assays.

4.12.2 Loci at which multiple genes were tested

As explained in section 4.5, multiple genes were studied at eight of the GWAS loci. Of these, three loci had no noteworthy results, three loci contained a gene which is a potential hit, and two loci showed a clear top hit. In three cases, the gene highlighted is not the one closest gene to the SNP.

Multi-gene loci with no noteworthy results

At the rs2287019 SNP, two genes were analysed: QPCTL (*QC, isoQC*) and GIPR (*Dh44-R1, Dh44-R2*). Only *isoQC* shows a score >0.7, but this is associated with mostly decrease phenotypes and so is not worth pursuing. Baranski et al. also found that neither of these genes affected TAG. *Dh44-R2* larvae do not show any difference in feeding behaviour between the fed and the fasted states.

Similarly, at the rs2241423 SNP locus, neither MAP2K5 (*dsor1*) nor LBXCOR1 (*fuss*) show a strong phenotype with neuronal RNAi. LBXCOR1 does show a noteworthy score with ubiquitous RNAi but this score is biased by a strong phenotype in just one assay so may a false positive result. Baranski et al. found no hits in this region.

TMEM18 (*Tmem18*), SH3YL1(*RtGEF*), and SNTG2 (*Syn2*) were phenotyped at the rs2867125 SNP locus, but none of these genes were hits based on either the score or the individual assay results. Baranski et al. also saw no TAG phenotypes in this region.

Multi-gene loci with potential hits

At the rs7359397 locus, the expression of genes correlates with the BMI SNP variant (Speliotes et al., 2010). In the screen here, only TUFM shows a phenotype worthy of further investigation, with increases in all assays for whole body RNAi. As explained earlier, the SH2B1 homologue *Lnk* was not tested here, but this gene is already known to affect the storage of carbohydrates and lipids (Slack et al., 2010; Song et al., 2010).

Gene	*Drosophila* orthologue	Homology (%)	Neuronal score	Whole body score
SH2B1	*Lnk*	-	not tested	not tested
APOB48R	-	-	-	-
SULT1A2	*St2*	29	0.67	lethal
AC138894.2	-	-	-	-
ATXN2L	*Atx2*	8	0.5	lethal
SBK1	*CG4945*	27	0.69	*tiptop*
	PKN	25	0.58	0.62
TUFM	*EfTuM*	60	0.67	0.87

At the rs713586 SNP. Only two of the genes had a *Drosophila* orthologue. One of these, RBJ (*rab21*), is a hit with both neuronal and ubiquitous RNAi.

Gene	*Drosophila* orthologue	Homology (%)	Neuronal score	Whole body score
RBJ	*rab21*	32	0.72	0.76
ADCY3	*ac3*	39	0.66	0.46
POMC	-	-	-	-
NCOA1	-	-	-	-

Three genes were studied at the rs2112347 locus. Of these, COL4A3BP (*cert*) may be worth further investigation as it has a score >0.7, but its individual assay results are complicated, with increases in feeding, but decreases in measures of physiology. Baranski et al. found no genes which affected TAG in this region.

Gene	*Drosophila* orthologue	Homology (%)	Neuronal score	Whole body score
FLJ35779	-	-	-	-
HMGCR	*hmgcr*	47	0.66	lethal
COL4A3BP	*cert*	44	0.78	0.59

Multi-gene loci with strong hits

The region near the rs2815752 SNP contains the gene NEGR1 as well as three other genes and a binding site for the transcription factor NKX6.1. There were no *Drosophila* homologues of LRRIQ3 and PTGER3, but ZRANB2 and NKX6.1 both have an orthologue and NEGR1 has three. All of the NEGR1 homologues have scores >0.7 which makes it a strong hit, whereas ZRANB2 and NKX6.1 do not. Further, *DIP-iota* mutant larvae behave in the same way regarding food irrespective of whether they are in the fed or the fasted state.

Gene	*Drosophila* orthologue	Homology (%)	Neuronal score	Whole body score
NEGR1	*lac*	24	0.73	male-lethal
	DIP-iota	26	0.86	male-lethal
	ama	24	0.77	semi-lethal
ZRANB2	*CG3732*	45	0.58	0.53
NKX6.1	*HGTX*	39	0.37	0.58
LRRIQ3	-	-	-	-
PTGER3	-	-	-	-

Two genes were investigated in relation to the rs10938397 SNP: GNPDA2 and GABRG1. Based on the *Drosophila* screen, GABRG1 is a more promising candidate for this locus, despite not being the nearest gene to the SNP: the orthologue *CG8916* shows a score >0.8 in the screen, and the orthologue *Grd* shows a strong increase in larval feeding behaviour.

Gene	*Drosophila* orthologue	Homology (%)	Neuronal score	Whole body score
GNPDA2	*Oscillin*	74	0.50	lethal
GABRG1	*CG8916*	32	0.81	0.75
	Grd	37	0.67	0.51

4.13 Published mammalian phenotypes of top hit genes

If a particular gene has already been studied in humans and/or mice, such data can be used to validate or discredit the results of the *Drosophila*-based screen used here. Conversely, if very little is known, then the gene is a prime candidate for further study to obtain novel results. Therefore, a PudMed search was performed for published data from mice and/or humans about the top genes from the screen.

4.13.1 Strong hits

The strong hits were genes which scored >0.8 in the *Drosophila*-based screen, or >0.7 with multiple orthologues, based on increase phenotypes. These genes were: PCSK1, NEGR1, MTIF3, TFAP2B, NRXN3, GABRG1, KLF9 and LRRN6C.

For three of these genes, mouse/human studies have shown an involvement in energy homeostasis. Firstly, PCSK1 has been identified to be the cause of some monogenic obesity disorders (see chapter 1) — mice/humans with KO mutations in PCSK1 show dramatically increased food intake, increased mass and hypoglycaemia (Dickinson et al., 2016; Zhu et al., 2002; Stijnen et al., 2016). Secondly, NEGR1 mutant mice have decreased body mass, food intake and physical activity (Lee et al., 2012; Dickinson et al., 2016; Speakman, 2013). Finally, mice with KO of MTIF3 show hypoactivity and increased body fat (Dickinson et al., 2016).

Although mice with KO of TFAP2B have been studied, they show a range of detrimental phenotypes across multiple organs and systems, which makes it difficult to identify any specific energy homeostasis phenotypes (Jackson Laboratory, 2018).

There were four genes for which no published studies were found related to energy homeostasis: NRXN3, GABRG1, KLF9 and LRRN6C.

An important point to note is that several of these genes have also been genetically linked to diseases that were used as negative controls in chapter 3, for example NRXN3 with autism and schizophrenia (Wang et al., 2018; Hu et al., 2013), LRRN6C with Parkinson's (Wu et al., 2011), and MTIF3 with Parkinson's (Behrouz et al., 2010).

4.13.2 Potential hits

The potential hits were genes which scored between 0.7 and 0.8 in the *Drosophila*-based screen with increase phenotypes. These genes were: CDKAL1, RBJ, COL4A3BP, BDNF, SEC16B, TUFM and NUDT3.

Again, there is published mammalian data which shows a role in energy homeostasis for some of these genes. CDKAL1 KO mice show significantly reduced fat storage, impaired insulin signalling and an increased respiratory exchange ratio, and this is amplified by HFD (Okamura et al., 2012; Dickinson et al., 2016). Mice with KO of RBJ also have an altered respiratory exchange ratio and glucose response (Dickinson et al., 2016). In humans, levels of RBJ in the plasma and adipose tissue are elevated in obese individuals and positively associated with leptin (Cherian et al., 2018).

For some of the genes, KO mutations have many detrimental phenotypes which makes the results difficult to interpret. For example, COL4A3BP mutations have been linked to severe intellectual disability in humans (Hamdan et al., 2014) and KO mice are very unhealthy (Wang et al., 2009; Dickinson et al., 2016). Similarly, mice with KO of BDNF have many developmental abnormalities, particularly in the brain (Jackson Laboratory, 2018), although there are many published papers exploring the function of BDNF in the nervous system.

Few phenotypes are published for SEC16B KO mice, but they have been shown to have altered cholesterol levels (Dickinson et al., 2016). There was no data specifically related to energy homeostasis for TUFM or NUDT3.

Again, some of these genes have been linked to the negative control diseases used in chapter 3 such as BDNF with schizophrenia (Jönsson et al., 2006).

4.13.3 Unhealthy genes

The unhealthy genes were those which scored >0.7 in the *Drosophila*-based screen, but had solely/mostly decrease phenotypes and thus were suspected of being unhealthy rather than having a specific lean phenotype. There were four such genes in the *Drosophila* screen, and published data for three of these suggests that mammals are similarly affected. Mice with KO of SLC39A8 are embryonic lethal (Dickinson et al., 2016) and two human patients with mutations in this gene showed developmental abnormalities, severe spasms and dwarfism (Park et al., 2015). Mice with KO of MTCH2 are embryonic lethal (Dickinson et al., 2016; Zaltsman et al., 2010). Similarly, mice with KO of RPL27A are homozygous lethal, and heterozygotes show retarded growth and ataxia (Terzian et al., 2011).

4.14 Discussion

In this chapter, the *Drosophila*-based screen which was validated in chapter 3 was used to explore genes which were identified by published GWASs to be related to human BMI.

4.14.1 Conservation of BMI GWAS genes

Human genes were selected for study which published GWASs have associated with variation in BMI. 79% of these human genes had at least one orthologue in *Drosophila* (section 4.5), which is higher than the genome-wide average of 53% (Wangler et al., 2017). Since human genes linked to monogenic disorders show increased levels of conservation in *Drosophila* (Fortini et al., 2000), this result in the GWAS genes is unsurprising.

UAS_G-siRNA lines were available to target all of the selected genes. Despite the fact that two of the lines died, the number of researchable genes was still 77%, which is high enough for a *Drosophila*-based screen to be a useful tool for exploration of the data.

Gene expression microarrays have shown that expression of genes at the GWAS BMI loci is enriched in the CNS (Locke et al., 2015). Published microarray data showed that 55 (out of 58) of the *Drosophila* genes are expressed in the larval, pupal and/or adult CNS (section 4.6). This conservation of expression patterns increases the likelihood of conservation of function of the genes (Wangler et al., 2017). Further, it shows that use of neuron-specific RNAi is a valid approach to study these genes. The three genes which are not expressed in the CNS showed no significant phenotype with neuronal RNAi, which increases confidence that the screen is not detecting false positive results. One caveat is that the microarray data is not 100% accurate, for example it shows no expression of *upd1* in the adult brain which contrasts to immunostaining data published by Beshel et al. (2017).

4.14.2 Lethality of BMI GWAS genes

59% of the LoF lines were lethal as adults (section 4.7), which is much higher than the genome-wide figure of 25% (Wangler et al., 2017). The true figure may be even higher because some transposable element insertions (the most common perturbation in the LoF lines used here) do not disrupt their target gene very strongly, dependent on the location of the insertion (Groth et al., 2004).

For the RNAi lines, 51% of the GD lines were lethal when crossed to actGAL4 (section 4.7). It is unsurprising that this figure is lower than the LoF lines, since RNAi only causes a

partial reduction in gene expression, but is still much higher than the genome average. The fifteen lethal *Drosophila* GD lines represent ten human genes, and viability data is available for eight of these in mice: six are lethal (Jackson Laboratory, 2018). This suggests that perhaps viability in *Drosophila* could be used as an indicator of whether a mouse LoF model will be viable, although the numbers here are too low to confidently draw such a conclusion. Four genes showed male-specific lethality in *Drosophila*, which tentatively suggests that if these genes are studied further in mice/humans it may be important to study both genders.

The high levels of lethality observed in the GWAS BMI genes suggests that they have a very important role within the body. Indeed, humans with loss-of-function mutations have not been identified for many of the GWAS BMI genes, and one possible explanation for this is that such mutations are so detrimental that humans are not viable (lab group internal discussion).

Half of the genes show the same viability/lethality in both the RNAi and LoF models, increasing confidence in these results. For the genes which were viable with RNAi but lethal with LoF, there are three possible explanations for the discrepancy:

1. RNAi reduces (rather than prevents) gene expression to levels of approximately 25% (Heigwer et al., 2018), and so there may be enough residual gene function to allow viability — this is likely the explanation for genes such as *isoQC* which showed a very unhealthy phenotype with whole body RNAi.

2. The siRNA was ineffective — this may explain why genes such as *Lrp1* and *Fancl* had the very lowest scores in the screen (i.e. they showed no phenotype).

3. The LoF perturbation also affected non-target genes, and it is these which are the cause of the lethal phenotype — for example the DNA sequence of *mEFTu1* overlaps with the *mip120* gene and so both genes could be affected by a genomic insertion, but the RNA exons do not overlap so could be differentially targeted by RNAi.

In contrast, for the RNAi lines which are viable with LoF but lethal with RNAi, there are also three possible explanations: (1) the siRNAs may have off-target effects; (2) the lethality of KK-background lines may be caused by over-expression of *tiptop* (see chapter 3.6.2), and/or (3) the location of the transposable element may mean that it does not significantly disrupt gene expression in the LoF line.

4.14.3 Phenotyping of GWAS BMI genes in *Drosophila*

Neuronal and whole-body RNAi were used to perturb expression of the 55 *Drosophila* genes selected for study (section 4.8). The resulting adult flies were subject to a suite of assays (validated in chapter 3) to assess energy homeostasis phenotypes of each gene: CAFE food intake, over-feeding dye, wet mass, glucose levels, climbing activity, and starvation resistance. In addition, LoF lines were used to measure larval feeding motivation, and adult feeding motivation was measured in the top RNAi hits.

Difficulties associated with *tiptop* over-expression

As was found in chapter 3.6.2, KK lines in which the UAS_G-siRNA transgene is inserted into the *tiptop* gene have a strong phenotype when crossed to elav-GAL4 in the screen. Therefore, it is important to compare such lines to the correct background in order to avoid false positive results.

For the KK lines which were lethal when crossed to act-GAL4, the background for neuronal RNAi was assigned to whichever gave the smaller phenotype, since it is better to have false negative results than false positives in this screen for candidate gene selection. If the larger background was chosen, then all 9 of the genes had a score that would be classed as a hit.

This method means that valuable true positive results are potentially falsely attributed to *tiptop* and therefore missed. Further, a phenotype may be present, but smaller than the effect of *tiptop* and therefore hidden. For both of these reasons, GD lines should be chosen in preference to KK lines whenever possible.

Comparison of GWAS BMI results to published literature

The results of the neuronal RNAi screen were compared to the recently published results from Baranski et al. (2018), who studied 17 fly genes in common (section 4.9). They found that 5 of these genes were associated with increased TAG storage in flies, and the matching results seen here for 4 of these genes provides independent validation of the accuracy of the screen. For the other gene, it is possible that their observed phenotype is due to an effect of the gene in the fat body — Baranski et al. used cgGAL4 to drive their RNAi in both the CNS and fat body, whereas elav-GAL4 was used here which only targets neurons.

Further, they found that 13% of their genes gave a lethal phenotype, whereas all of our crosses were viable with neuronal RNAi. This may be because of the wider expression of their RNAi (using cg-GAL4), and/or because they chose KK lines wherever possible but did not mention a control for *tiptop* over-expression. Therefore, the screen used here is likely to be able to generate results for a higher proportion of genes.

Finally, for 3 of the genes, Baranski et al. did not see any TAG phenotype, but because multiple assays were used here, the screen was able to detect some more subtle phenotypes of energy homeostasis.

4.14.4 Interpretation of results according to human gene

The aim of this screen was to select genes for further study from a list of human SNPs that were identified by GWASs of human BMI. In summary, 36 genetic loci were studied. 7 of the BMI loci could not be interrogated with the *Drosophila* screen because the constituent genes did not have a fly orthologue, or the siRNA lines died.

Single gene loci

At 19 loci, only the gene closest to the SNP was studied. Of these, 11 genes were not a hit in the *Drosophila*-based screen. In some cases, this was due to potential sickness, for example RPL27A. In other cases, no phenotype was detected in the screen assays, and there are several possible explanations for this:

1. The siRNAs were not effective, either because of their sequence or because their genomic location means that they are only expressed at low levels (Heigwer et al., 2018).

2. The *Drosophila* gene may not be a functional orthologue of the human gene due to differences in mammalian and insect physiology. (Wangler et al., 2017)

3. Some genes may only cause a phenotype when up-regulated, and so would not be detected by the RNAi approach used here.

4. The screen was designed to avoid false positive results rather than false negatives. For example, genes were considered hits if their score was >0.7, but this cut-off was based on the results of the negative control genes in chapter 3. As discussed previously, these negative control genes are not necessarily independent of energy homeostasis, and so

perhaps the stringency of the screen was too high. Indeed, as mentioned in section 4.13, there is data linking many of the GWAS BMI genes to the same negative control diseases.

5. The screen was designed to focus on energy intake and physiology, rather than on energy expenditure. Therefore genes which affect the latter may not be detected. For example, CADM2 was not a hit in the screen, but CADM2 KO mice have been shown to have altered energy expenditure (Yan et al., 2018), but there is currently no published data on food intake.

6. For the *Drosophila* lines which were potentially affected by *tiptop*, the background chosen was that which gave the smaller phenotype. If the incorrect background was chosen, hits would be overlooked.

7. The human gene does not actually function in energy homeostasis. For such genes, it may be useful to explore other genes within the region. For example, Baranski et al. (2018) looked at all genes within a 500kb window of each SNP.

Since the purpose of this high-throughput screen was to find top candidate genes involved in energy homeostasis, and not to provide a definitive list of the function of the individual genes, false negative results are not problematic. Instead, they increase the confidence in those genes which were deemed hits.

At the single gene loci, five genes were strong hits with neuronal RNAi: TFAP2B (*TfAP-2*), NRXN3 (*Nrx-1*), PCSK1 (*fur1, rab2*), KLF9 (*sp1, cabot*), and LRRN6C (*caps, CG14672*). BDNF (*NT1*) and CDKAL1 (*CG6550*) were modest hits with ubiquitous RNAi, NUDT3 (*Aps*) was a hit with neuronal RNAi, and SEC16B (*dSec16*) was a hit with both. Overall, more hits were obtained with neuronal RNAi than with ubiquitous RNAi, which is unsurprising given the role that the CNS plays in the control of energy homeostasis (see chapter 1), and the fact that whole-body genetic perturbations are often detrimental to an organism (Yazdi et al., 2015).

Multi-gene loci

At eight loci, multiple genes were studied. Two genes were identified as top hits worthy of further study (NEGR1 and GABRG1) and two as potential hits (COL4A3BP, RBJ). Half of these (GABRG1 and COL4A3BP) are not the gene closest to the SNP, highlighting the utility of using a high-throughput organism to perform an unbiased screen for phenotypes.

The other four loci showed no noteworthy results, and this is likely due to the same reasons discussed above in relation to the single gene loci.

Genes with multiple orthologues

For a few genes (PCSK1, LRRN6C, NEGR1, KLF9), more than one *Drosophila* orthologue was studied and found to be a hit, which increases confidence that the corresponding human gene is worth further study. For other genes (GABRG1, MTCH2), multiple orthologues were phenotyped but only one was found to be a hit. For both of these reasons, it is perhaps worth testing multiple homologues (where possible) in future screens. This depends on a balance between the impact of false negative results versus the size of, and the time available for, the screen.

Published data for the top hit genes

The *Drosophila*-based screen performed in this chapter provided data to support more detailed experiments on the mechanism of action for these genes in influencing human BMI.

Based on a search of published literature (section 4.13), several of the top hit genes do have an energy homeostasis phenotype in mammalian models (PCSK1, NEGR1, MTIF3, CDKAL1, BDNF), which provides evidence that the screen is producing accurate and relevant results. This is also true for the genes which produced lean/unhealthy phenotypes in the *Drosophila* screen (SLC39A8, MTCH2, RPL27A).

On the other hand, some genes (LRRN6C, KLF9, NRXN3, GABRG1, SEC16B, TUFM, NUDT3, COL4A3BP, TFAP2B) are relatively unexplored in relation to energy homeostasis, marking them as prime candidates for further study.

4.14.5 Reliability and limitations of the results

The use of multiple assays in the screen and a stringent cut-off score together reduced the chance of results being false positives. Therefore, it is likely that the results observed are a true representation of the phenotype of the line tested. However, it is possible that some of the siRNA lines used had off-target effects which were responsible for the phenotype observed. To confirm that the results were indeed due to the gene of interest, several different models could be used. Firstly, some genes (although not all) have more than one different siRNA line available which could be used to repeat the experiment. Secondly, LoF *Drosophila* lines could

be obtained and studied, although again these do not exist for all genes, and these whole-body mutations may not be able to replicate a neuron-specific perturbation. Alternatively, for some genes, there is relevant data published which could be compared. However, none of these methods are applicable to all genes, and all have draw-backs. Therefore, in chapter 5, methods of generating precise perturbations of the genes of interest *de novo* was explored. Since the initial list of genes are already suspected of playing a role in energy homeostasis, it could perhaps be argued that the risk of a hit in the screen being a false positive is low.

The main limitation of the screen is that not all genes can be studied: not all human genes have an identifiable homologue in *Drosophila*; some of the genes do not have a healthy siRNA line that produces efficient knock-down; and it is difficult to be certain of phenotypes in the *tiptop*-affected lines. Further, the *Drosophila* genes were chosen based on sequence homology to the human genes, and it is possible that they are not also functional homologues.

4.14.6 Further work

There are several experiments which could be performed based on the results in this chapter. Firstly, the results for the top candidate genes could be replicated in a different model in order to check for off-target effects (a method for doing so is explored in chapter 5).

Secondly, for those loci at which the gene(s) studied were not hits in the *Drosophila*-based screen, it may be beneficial to screen more of the genes in the region. Perhaps initially just the top hits from Baranski et al.'s results (who examined all genes within 500kb of each SNP) could be studied.

Further, at the time that this project began, 32 loci had been identified that are associated with BMI (Lu and Loos, 2013). Now, more than 100 such loci are known (Locke et al., 2015; Akiyama et al., 2017), and this number will likely increase as technological and computational techniques improve even further. Experiments in mice cannot keep pace, and thus a screen using *Drosophila* would be beneficial to decide which of the new candidate genes are worth pursuing.

Finally, the top hits identified here (PCSK1, NEGR1, MTIF3, CDKAL1, BDNF, LRRN6C, KLF9, NRXN3, GABRG1, SEC16B, TUFM, NUDT3, COL4A3BP, TFAP2B) should be investigated in complex model organisms such as mice, particularly those genes for which not much is currently known. This could involve looking at where the genes are expressed, what conditions alter their expression, and what happens to a mouse if expression is deliberately perturbed. The detrimental effects seen in many of the LoF and ubiquitous RNAi lines here

suggests that, even though tissue-specific experiments are not as easy in mice as they are in *Drosophila*, it is worth investing in such experiments.

4.15 Summary

In conclusion, in this chapter the *Drosophila*-based screen that was developed in chapter 3 was used to study a series of human genes identified by GWAS to influence BMI. *Drosophila* homologues were identified of these genes, and shown to be (mostly) expressed in the CNS and enriched for lethal phenotypes. These were phenotyped using a suite of energy homeostasis assays, allowing selection of top candidate genes for further study. Published results already exist for some of these genes, and many showed the expected phenotypes based on the *Drosophila* results. Others have not been studied making them favourable for investigation. Many of the positive findings from the fly screen implicate the nearest gene to the SNP, but there are also some loci for which a more distal gene is the top candidate. Thus, the results in this chapter provide data to support more detailed experiments in mammals on the mechanism of action for certain genes in relation to human BMI.

Chapter 5

CRISPR gene editing in *Drosophila*

5.1 The specificity of RNAi and LoF in *Drosophila*

As discussed in chapters 1 and 3, one of the main advantages of using *Drosophila* as a model organism is the availability of large stocks which can be used to genetically perturb almost any gene of interest. These stocks are convenient, especially for large-scale screens, but have drawbacks (as discussed in chapter 3) in particular with regards to specificity. For LoF lines, the genetic perturbation can delete or disrupt multiple genes and regulatory elements (Groth et al., 2004).

The screen of GWAS BMI genes carried out in chapter 4 used flies (obtained from VDRC: www.vdrc.at) which express siRNAs that target the genes of interest. But siRNAs can have off-target effects if they have enough sequence similarity to hybridise with non-target RNAs (Seinen et al., 2011). Of the 142 siRNA lines used in this thesis (chapters 3, 4 and 6), 24 lines are predicted by VDRC to have one off-target effect and three lines are predicted to have at least two.

One way to check that a result obtained is not due to off-target effects is to replicate the results using a different line. However, only 60% of *Drosophila* genes have an additional UAS_G-siRNA line available at VDRC to use for comparison. Therefore, in this chapter, a method of generating *de novo* mutations in a specific target gene was explored for use in confirming results from an RNAi-based *Drosophila* screen.

5.2 Gene editing using CRISPR

CRISPR is a recently developed technique for genetic engineering. The endogenous function of CRISPR-Cas9 as a bacterial immune system was discovered in the 1990s, and in 2013 two papers were published showing that this system could be adapted to use in genetic editing of eukaryotic cells (Cong et al., 2013; Mali et al., 2013). The first use of CRISPR in *Drosophila* followed soon after (Gratz et al., 2013).

The basic principle of generating mutations using CRISPR is to introduce two components into the target cells: Cas9 protein and guide (g)RNAs. Cas9 is an endonuclease which creates double strand breaks (DSBs) with blunt ends in DNA. gRNAs are approximately 100 nucleotides long and have a specific secondary structure which binds and activates Cas9, as well as a 20 nucleotide sequence which is homologous to the target DNA. Cas9 is guided to a target sequence in the cell's DNA by the gRNA's sequence homology. The homologous sequence in the gRNA can be altered to target Cas9 to different genes of interest, where Cas9 can then cleave the DNA. Repair of the DSB by non-homologous end-joining (NHEJ) is usually imperfect and creates indel mutations, so disrupting the target gene.

One limitation of CRISPR is that Cas9 requires the presence of a protospacer-associated motif (PAM) sequence (NGG) at the 3' end of the target sequence, restricting the possible target sites. Whilst it is possible that a desired sequence can not be targeted, this is not usually a problem for creating indel mutations because it is rare that the mutation must be in a very precise location.

Cas9 protein is unlikely to act using endogenous RNAs (which would cause non-specific effects) because they lack the Cas9-activating secondary structure of the gRNA. However, gRNAs can have off-target effects because the homology region between gRNAs and genomic DNA is small (only 20bp) and some base mismatches are tolerated (Hsu et al., 2013; Fu et al., 2013; Pattanayak et al., 2013). Since it is unlikely that the off-targets of CRISPR would be the same as that of RNAi, if the same phenotype is seen with both models then it is likely to be a true phenotype of the gene of interest.

5.3 Chapter aims

In this chapter, three methods of CRISPR were explored for their potential to use as part of a *Drosophila*-based screen of genes involved in the regulation of human energy homeostasis.

Firstly, creation of lines with uniform whole-body mutation of the target genes was explored in section 5.6. This method is referred to as homogeneous CRISPR.

Secondly, a method to create tissue-specific but non-homogeneous mutations in fly somatic cells was explored in section 5.7. This method is referred to as tissue-specific CRISPR.

Finally, a method similar to tissue-specific CRISPR but mediated by the UAS_G-GAL4 system was investigated in section 5.8 and is referred to as GAL4 CRISPR.

Results

5.4 Selection of target genes

To test methods of modifying genes by CRISPR, two target genes were chosen: *fuss* and *rab21* which are homologues of the human genes LBXCOR1 and RBJ respectively. These two genes were chosen because (i) they are viable as actin knock-downs, increasing the chance that CRISPR mutants will be viable, and (ii) results from preliminary RNAi data were promising. Another reason for using *fuss* is that its genomic location makes it difficult to target and thus good for testing the system - *fuss* is located on chromosome 4 which is largely heterochromatic, relatively unstudied, and difficult to obtain balancer chromosomes for (Sun et al., 2000).

Once RNAi data collection was complete, *rab21* had a score of 0.72 for neuronal RNAi and 0.76 for ubiquitous RNAi and was considered a hit (see chapter 4). *fuss* had a score of 0.54 for neuronal RNAi and 0.74 for ubiquitous RNAi but was not considered a hit (see chapter 4).

5.5 Design of gRNAs

In order to increase the likelihood of gene deletion, each gene was targeted with two gRNAs. Using this method, deletions of 6.1kb of the intervening DNA have been demonstrated in *Drosophila* (Gratz et al., 2013) and even larger deletions have been shown in other model systems including 65kb in mice (Zhang et al., 2015a). Possible gRNA sequences were identified using the CRISPOR tool (Haeussler et al., 2016) and two chosen for each gene based on the following criteria:

- 20bp long — this is widely reported in the literature to be a good length to balance the specificity and efficiency of CRISPR (for example see Zhang et al. (2016))
- Predicted to have a high efficiency of mutation (Fusi-Doench score >50, Doench et al. (2016))
- No predicted off-targets with three or fewer mismatches, to ensure specificity
- Position of the gRNA target site within the gene of interest:
  - One gRNA near the beginning of the coding sequence and one near the end to increase the likelihood of deletion of key regions(s) of the gene
  - Regions of the gene which were present in all isoforms
  - The *fuss* gene overlaps with the *sphinx* gene and thus the gRNAs were chosen to avoid *sphinx*. *fuss* also overlaps with a non-coding RNA, but this was not annotated at the time the gRNAs were designed

The location of the gRNAs relative to each gene are shown in figures 5.1 and 5.2 and the sequences are given in methods section 2.5.

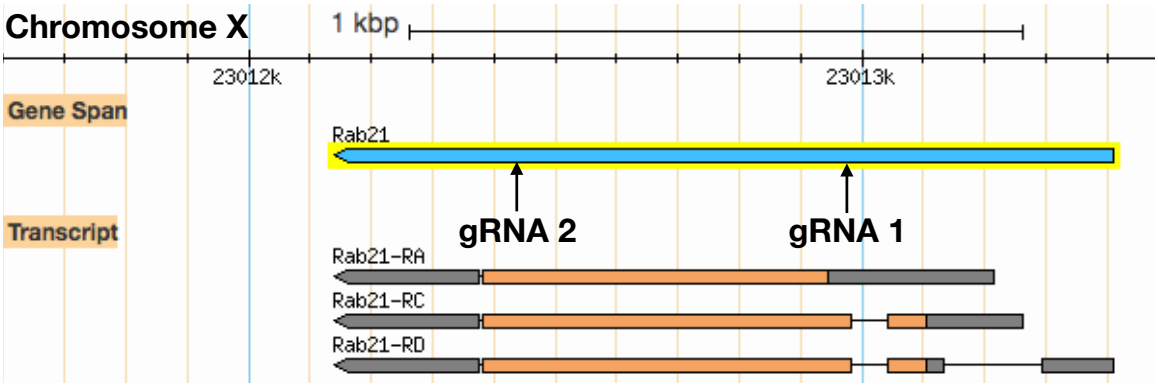


Fig. 5.1 Map of the *rab21* gene with locations of the designed gRNAs. The gene is shown in blue, introns as thin lines, coding sequence exons in orange, and UTRs in grey.

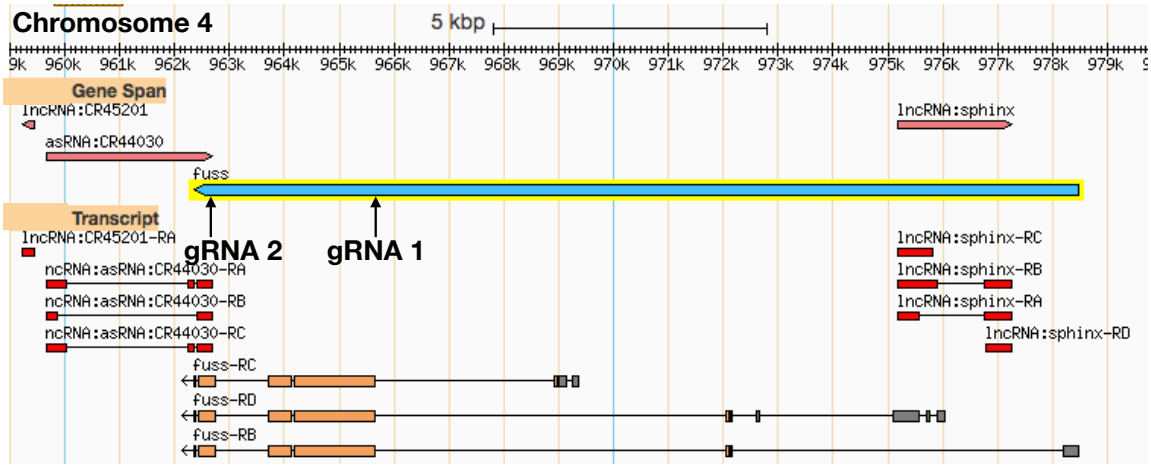
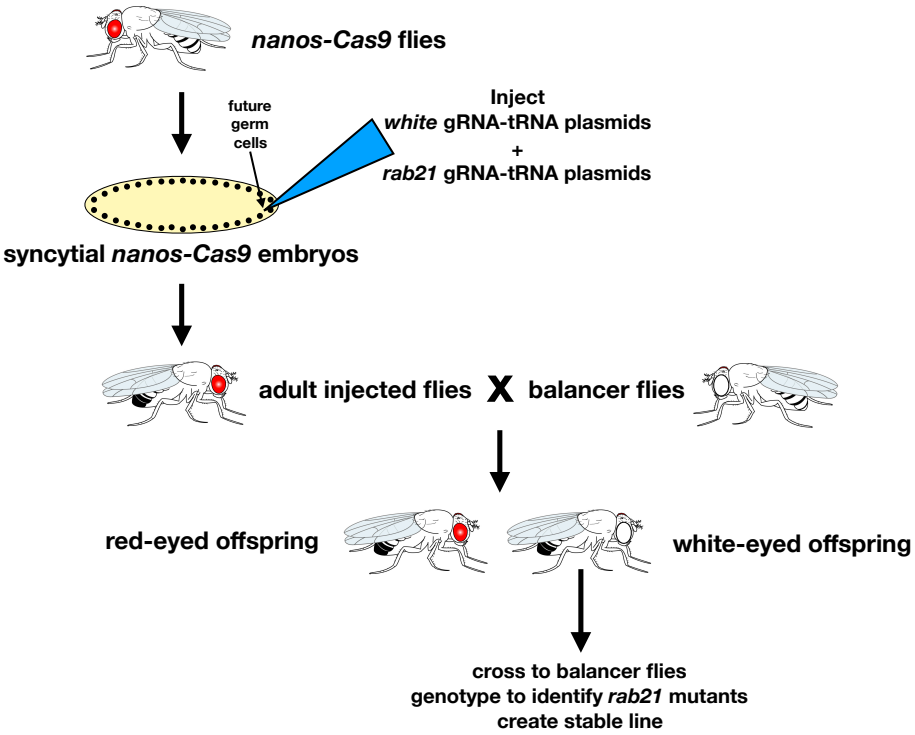


Fig. 5.2 Map of the *fuss* gene with locations of the designed gRNAs and overlapping genes. The gene is shown in blue, introns as thin lines, coding sequence exons in orange, and UTRs in grey.

5.6 Mutation of genes by homogeneous CRISPR

5.6.1 Homogeneous CRISPR strategy

The overall strategy to generate homogeneous mutants by CRISPR is summarised in the following figure which uses *rab21* as an example:



Embryo injection

To produce a fly with the same mutation throughout the body (non-mosaic), the germ cells of the parents need to be targeted. Therefore, the gRNA mixtures were injected into dechorionated syncytial preblastoderm *Drosophila* embryos before formation of the germ cells. In the injected *Drosophila* line, Cas9 expression is controlled by the *nanos* promoter and so is only present in the area of the pole cells. The injected embryos were grown to adults, crossed to balancer flies, and then future generations examined for mutation of the target genes.

Injection plasmid design

Port *et al.* have previously reported creation of a vector (pCFD5) in which two gRNAs are separated by tRNAs (Port and Bullock, 2016). When in a cell, the U6:3 promoter present in the plasmid drives ubiquitous transcription by RNA polymerase II/III of the plasmid, so producing a single transcript of gRNAs separated by tRNAs (figure 5.3). The endogenous tRNA processing machinery liberates tRNAs from this transcript with concomitant release of the gRNAs.

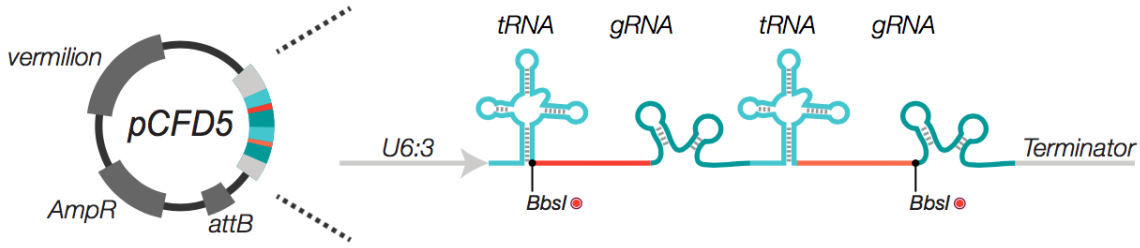


Fig. 5.3 Map of plasmid pCFD5 which was used to introduce gRNAs into *Drosophila*. Diagram obtained from www.crisprflydesign.org/plasmids/

Using genetic cloning techniques described in section 2.5, a single plasmid for each target gene was created containing both of the designed gRNAs. The gRNA-tRNA sequences designed to target *fuss* and *rab21* were correctly created and inserted into the pCFD5 plasmid, as confirmed by sequencing.

Screening by co-mutation of *white*

To facilitate screening for mutants, a plasmid with gRNAs that target the *white* gene was co-injected with the gRNA plasmids targeting *fuss* or *rab21*. In its wild-type form, the *white* gene produces the red colour of *Drosophila* eyes. KO mutation of the *white* gene results in flies with white eyes. Mutation of *white* can be assessed quickly and simply by looking at eye colour using a bench microscope.

Using this approach Ge et al. (2016) showed that flies which gain a mutated *white* gene (as evidenced by white eyes) have a higher rate of mutation of the gene of interest compared to flies which retain wild-type *white* (as evidenced by red eyes). Selecting flies with white eyes reduces the number of flies that need to be screened molecularly which is a time-consuming process. gRNA sequences to target *white* were taken from the literature (Ge et al., 2016). A tRNA-gRNA plasmid targeting the *white* gene was created using the same method as above and confirmed by sequencing.

5.6.2 Production of mutant flies by homogeneous CRISPR

rab21

202 nanos-Cas9 embryos were injected with gRNA-tRNA plasmids targeting *rab21* and *white*. 41 (20%) of the injected embryos survived to adulthood and 7 (3%) of these produced some progeny with white eyes (i.e. a mutated *white* gene). 4-10 offspring from each of the 7 white-eyed fly lines were screened using PCR to look for deletion of the DNA between the 2 gRNA sites. In wild-type flies the PCR product is 785bp long (for example see figure 5.4), but in mutants the expected product would be approximately 340bp depending on the exact site of the mutation. No *rab21* mutants were found using PCR screening.

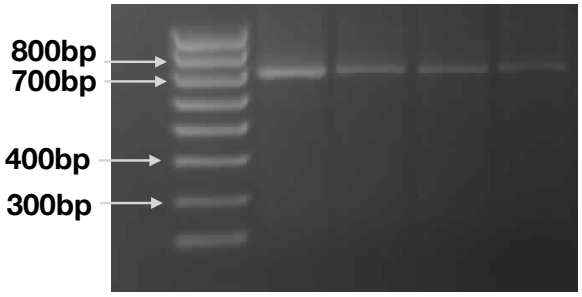


Fig. 5.4 PCR spanning the CRISPR target site did not reveal any large deletion mutations in offspring of the gRNA injected flies.

Therefore the *rab21* gene was sequenced in these flies. This revealed mutations in flies derived from 3 (1%) of the injected embryos. All of these mutations were at the gRNA site closer to the start of the gene, with none at the second site.

In the first mutant line, 4 out of 10 of the progeny contained a 13bp deletion in *rab21*, as shown in figure 5.5. This causes a frameshift which is predicted to result in 22 or 55 amino acids (depending on isoform) of wild-type sequence, followed by 5 incorrect amino acids and then a stop codon before the GTPase domain. For comparison, the wild type protein is 189 or 222 amino acids long. This line is referred to as $rab21^{\Delta 13}$ henceforth.

In the second line, 3 out of 4 of the progeny had a deletion of 6bp in *rab21* (figure 5.6). This results in deletion of only 2 amino acids from the protein. This line is referred to as $rab21^{\Delta 6}$.

In the final line, there was a deletion of 9bp and an addition of 3bp in *rab21* in 2 out of 5 of the flies, as shown in figure 5.7. The resulting proteins have 1 incorrect amino acid and 2 missing amino acids at the target site. Unfortunately, the founder fly died and so this mutation could not be tested further.

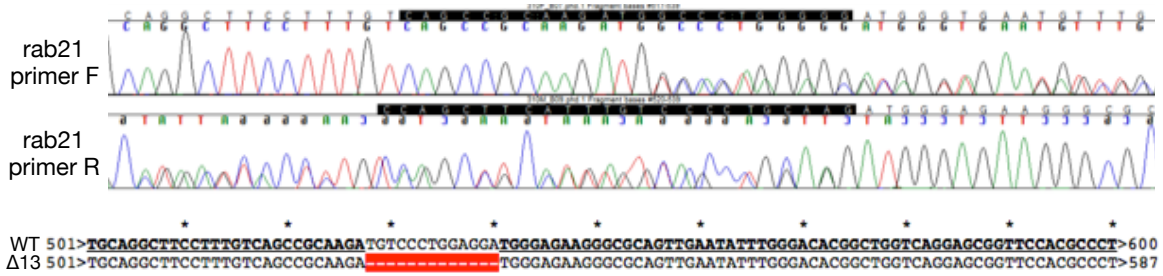


Fig. 5.5 Sequencing showed deletion of 13bp at the gRNA1 target site in *rab21* in the *rab21*$^{\Delta 13}$ line of injected *Drosophila*. See figure 5.1 for the location of gRNA1 within *rab21*.

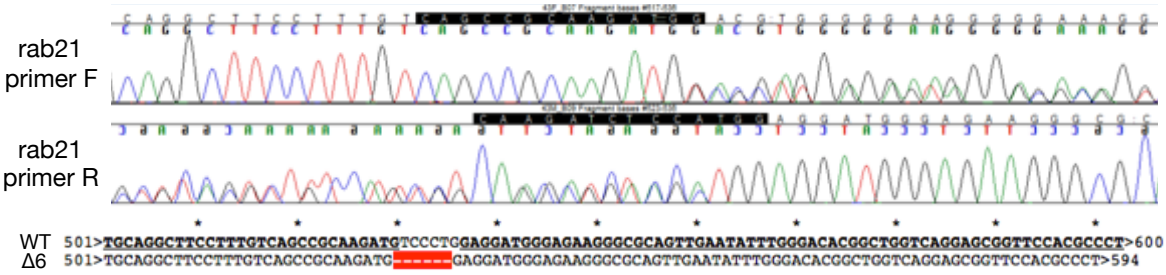


Fig. 5.6 Sequencing showed deletion of 6bp at the gRNA1 target site in *rab21* in the *rab21*$^{\Delta 6}$ line of injected *Drosophila*. See figure 5.1 for the location of gRNA1 within *rab21*.

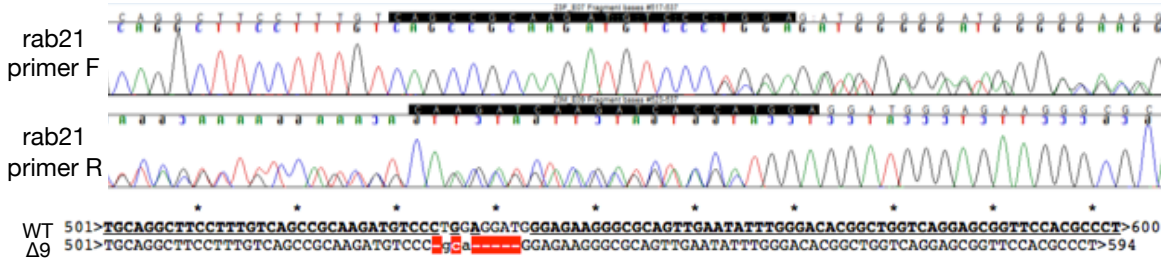


Fig. 5.7 A line of injected *Drosophila* had a 9bp deletion and 3bp insertion at the gRNA1 target site in *rab21*. See figure 5.1 for the location of gRNA1 within *rab21*.

One of the white-eyed flies which did not have any mutations in *rab21* was selected to use as a background-matched control. This line is referred to as $rab21^{WT}$ henceforth.

fuss

191 nanos-Cas9 embryos were injected with gRNA-tRNA plasmids targeting *fuss* and *white*. Only 17 (9%) of the injected embryos survived to adulthood, and none of these produced progeny with white eyes. A small number of the red-eyed flies were genotyped by PCR but none showed deletion of the DNA between the two gRNA target sites.

white

The *white* gene in some of the mutant *rab21* flies was also sequenced. Although the exact mutations varied between lines, there were many examples of deletion of the DNA in between the two gRNA sites. For example, figure 5.8 shows a deletion of the 209bp from the middle of the first gRNA to the middle of the second gRNA in one of the white-eyed flies.

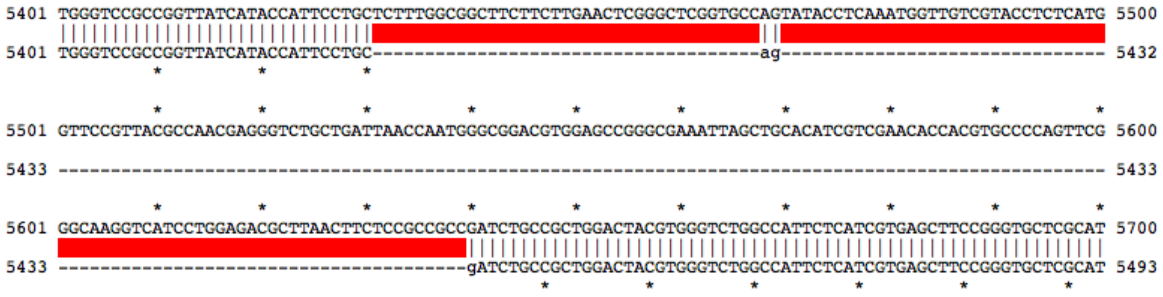


Fig. 5.8 Sequencing showed deletion of 209bp in the *white* gene in one line of injected *Drosophila* (bottom) compared to wild-type (top)

5.6.3 Efficiency of disruption of *rab21* expression in mutant flies

The two mutant lines, $rab21^{\Delta 6}$ and $rab21^{\Delta 13}$, were tested to see whether their mutations affected expression of the *rab21* gene. qPCR showed that RNA levels of *rab21* were decreased in the $rab21^{\Delta 13}$ mutants (figure 5.9). For the $rab21^{\Delta 6}$ mutants, there was a trend towards a decrease but it did not reach significance. Even in the $rab21^{\Delta 13}$ line, the RNA levels were not as strongly affected as they were by ubiquitous RNAi (figure 5.10).

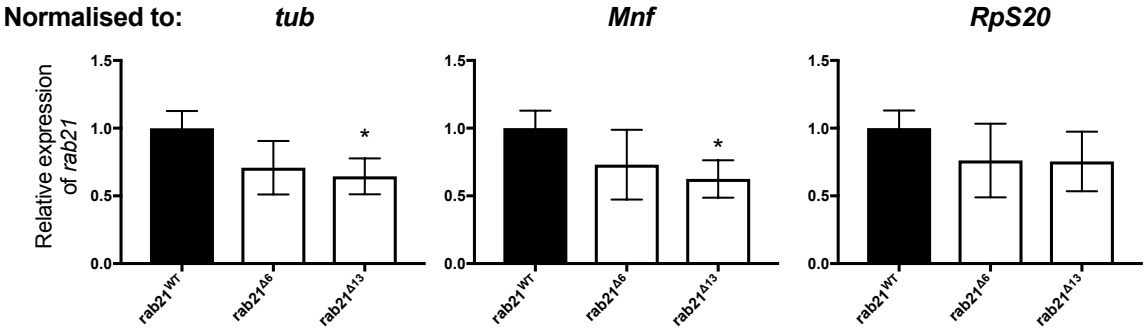


Fig. 5.9 Levels of *rab21* RNA are decreased in the homogeneous CRISPR mutant flies, as shown by qPCR. Mean ± SEM is plotted of 5 repeats each containing 10 flies. Results were compared to $rab21^{WT}$ by Student's t-test. * p<0.05

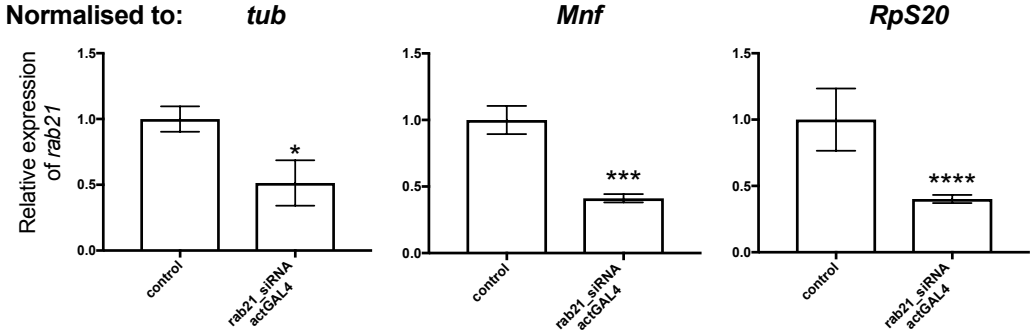


Fig. 5.10 RNA levels of *rab21* are decreased by actGAL4-driven RNAi. Mean ±SEM is plotted of 5 repeats each containing 10 flies. Results were compared by Student's t-test.
* p<0.05, *** p<0.001, **** p<0.0001

5.6.4 Phenotype of *rab21* mutant flies

It was immediately apparent that the *rab21* mutant flies were less fecund than wildtype, and the $rab21^{\Delta 13}$ flies were more affected than the $rab21^{\Delta 6}$ flies (figure 5.11h). Fecundity was not measured in the RNAi flies.

In addition, the phenotype (in the same screening assays used in chapter 4) of flies with ubiquitous RNAi of *rab21* was compared to that of flies with CRISPR-induced mutation of *rab21*. The results (figure 5.11) are broadly in agreement: there was a significant increase in dye food over-feeding and no phenotype for mass or CAFE feeding. Two results for which there was a trend with RNAi (a decrease in climbing and a decrease in starvation resistance) became significant with CRISPR. The only disagreement in the results is the glucose assay, where no phenotype is seen for CRISPR but a decrease is seen for RNAi. Finally, it was possible to measure TAG in the CRISPR flies but levels were undetectably low in the RNAi flies (see chapter 3.4).

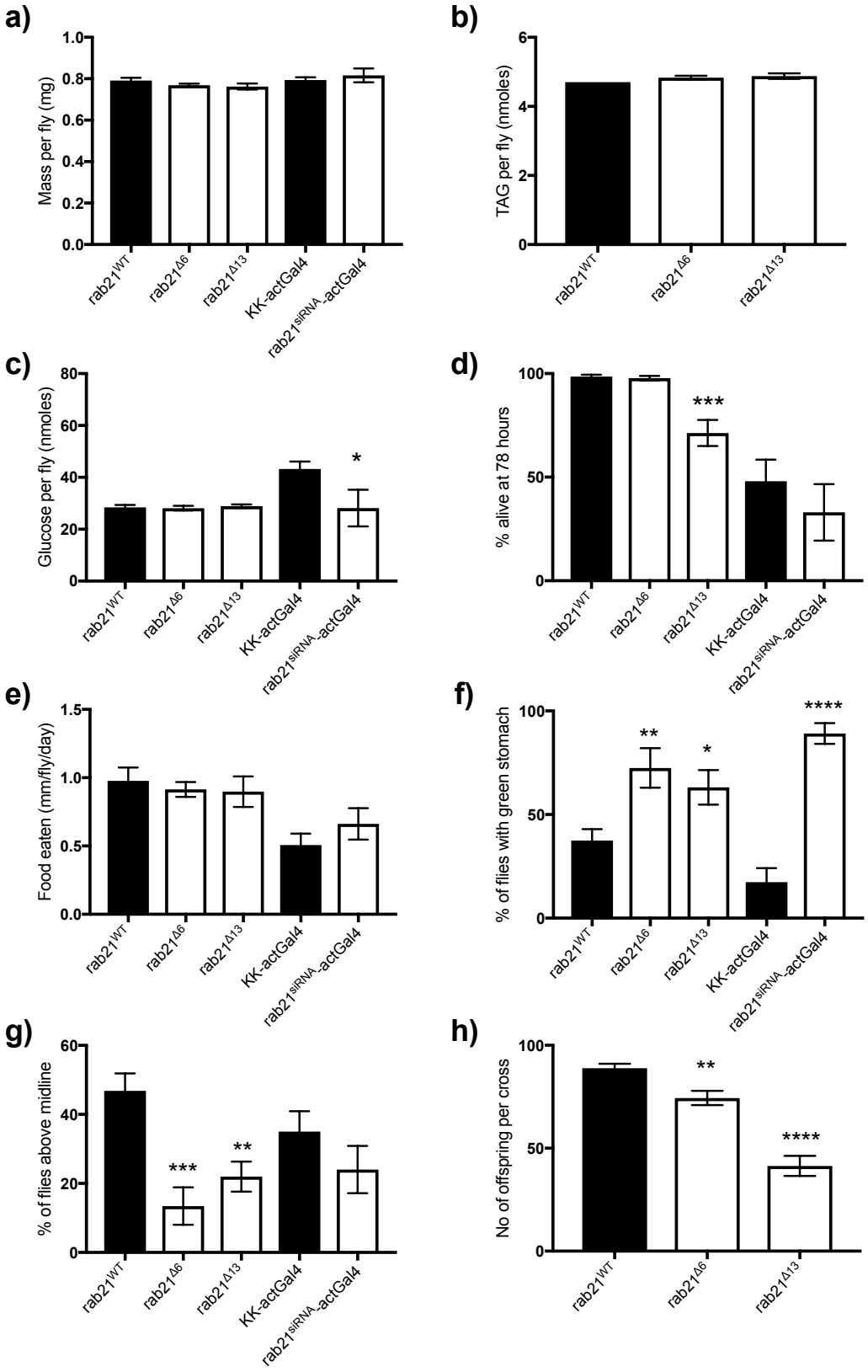


Fig. 5.11 *rab21* mutants generated by homogeneous CRISPR show a similar phenotype to flies with whole body RNAi of *rab21*. Background control lines are black and mutant/RNAi lines are white. Mean ± SEM is plotted of 5 repeats. Results were compared to appropriate background by Student's t-test. * $p<0.05$, ** $p<0.01$, ***$p<0.001$

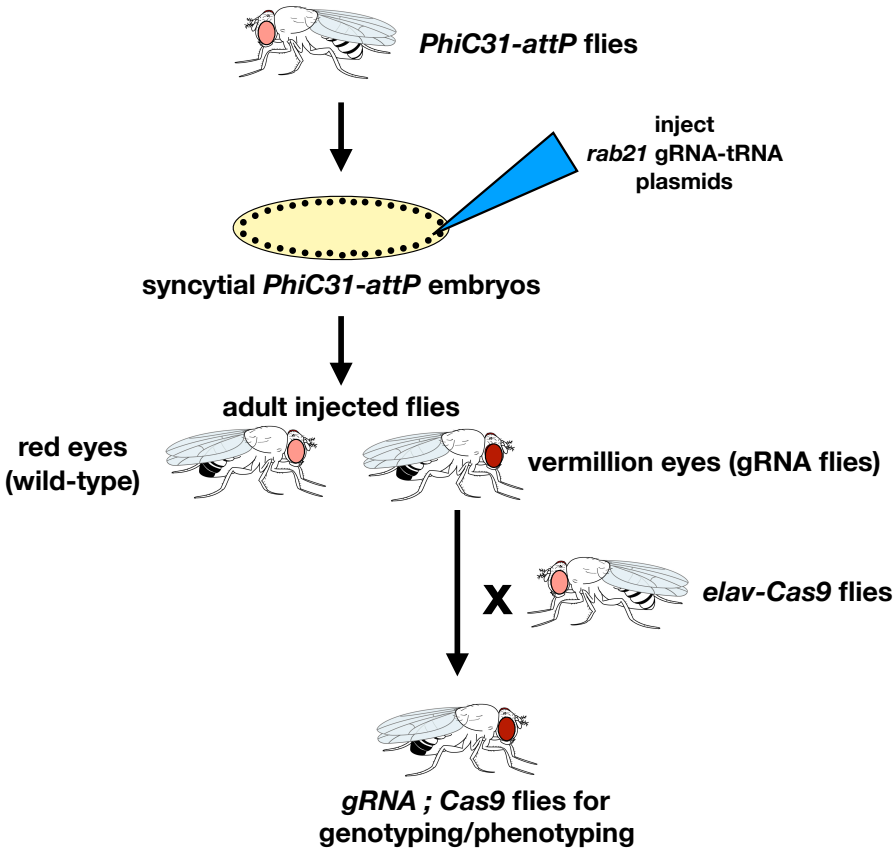
5.7 Tissue-specific CRISPR

The screen of GWAS genes carried out in chapter 4 focussed mainly on neuronal RNAi, and so results from a line generated by homogeneous CRISPR (in which the gene is KO throughout the whole body) may not always be relevant for comparison. Further, as discussed in chapter 3, a whole body KO mutation is detrimental or lethal for many genes which limits the information available from the homogeneous CRISPR model.

Therefore, to maximise the usefulness of creating *de novo* CRISPR mutations for the screen, the feasibility of neuron-specific CRISPR was explored.

5.7.1 Strategy

The overall strategy to generate tissue-specific mutants by CRISPR is summarised in the following figure:



Embryo injection

The gRNA-tRNA plasmids created in section 5.6 contained an attB site. Thus, the rab21-gRNA and fuss-gRNA plasmids were separately injected into *Drosophila* embryos which expressed PhiC31 integrase under the control of the *nanos* promoter. PhiC31 integrase promotes sequence-specific recombination between the plasmid attB site and a genomic attP site (in this case on chromosome 2), thus integrating the plasmid into the *Drosophila* genome. The PhiC31 integrase was under the control of the *nanos* promoter so, as before, the genomic integration should only occur in the germ cells.

Screening for plasmid integration

The injected embryos were grown to adulthood and their offspring screened for plasmid integration. The gRNA plasmids also contained the *vermillion* marker gene sequence and so it was therefore possible to screen the flies for plasmid insertion simply by the presence of vermillion-coloured eyes. For both target genes, many flies had successful integration of the plasmid but only one founder fly was used to create a line in order to ensure homogeneity. These lines are referred to henceforth as rab21-gRNA flies and fuss-gRNA flies.

Introduction of Cas9

As before, the gRNAs are downstream of a U6:3 promoter and so are ubiquitously expressed throughout the fly. When crossed to a *Drosophila* line expressing Cas9, CRISPR-induced mutation of the target gene can occur in the offspring. Several lines are available with Cas9 under the control of different promoters which affect when/where Cas9 is expressed and thus when/where mutation can occur. Mutations created in this way are somatic (unless a germ-line Cas9 is used) and are likely to be different in each cell (mosaic).

5.7.2 Gene expression after tissue-specific CRIPSR

Although the aim was to create neuron-specific mutations, a Cas9 line with an appropriate promoter was not immediately available for doing so. Further, it is easier to observe the effectiveness of the method (gene expression levels) using whole-body mutation because no tissue dissection is required. Therefore, whilst waiting for delivery of an elav-Cas9 line, the efficiency of the tissue-specific CRISPR method was first tested by crossing the gRNA flies to a line with ubiquitously expressed Cas9 (actin5C-Cas9)

When the rab21-gRNA line was crossed to actin5C-Cas9, there was no change in the levels of *rab21* RNA compared to either parental line as shown by qPCR (figure 5.12). This is in comparison to both ubiquitous RNAi and homogeneous CRISPR which caused a decrease in expression of *rab21* shown previously (figures 5.9 and 5.10).

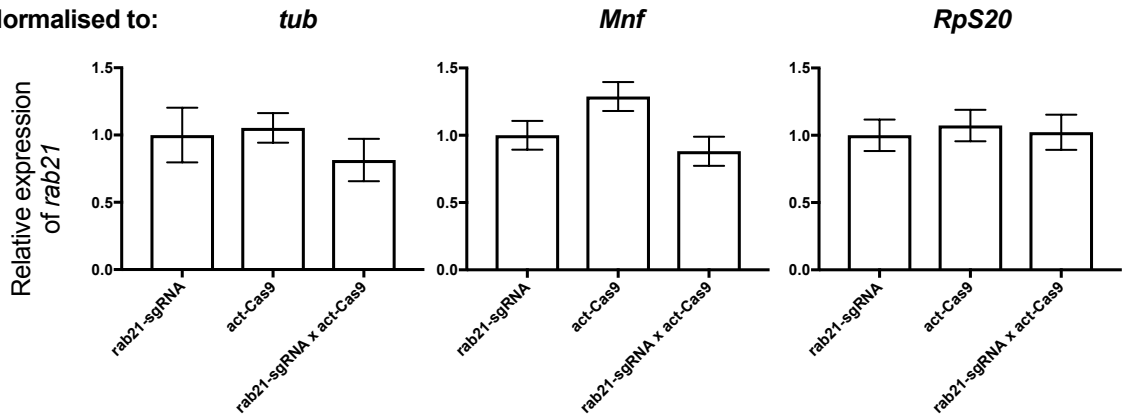


Fig. 5.12 No changes in *rab21* expression were detected after tissue-specific CRISPR. Mean ± SEM is plotted of 6 repeats each containing 10 flies. The results of each genotype were compared by Student's t-test.

When the fuss-gRNA line was crossed to actin5C-Cas9, qPCR showed an increase in the level of *fuss* RNA (figure 5.13). By comparison, ubiquitous RNAi caused a decrease in expression of *fuss* (figure 5.14).

Since it was not possible to show effective perturbation of *rab21* or *fuss* expression, the lines were not phenotyped.

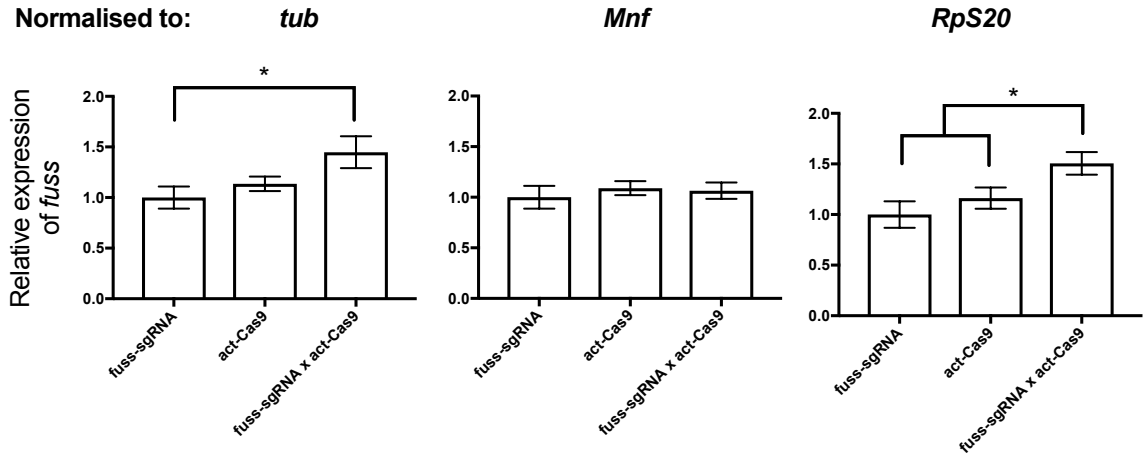


Fig. 5.13 An increase in expression of *fuss* was detected after tissue-specific CRISPR. Mean ± SEM is plotted of 6 repeats each containing 10 flies. The results of each genotype were compared by Student's t-test. * p<0.05

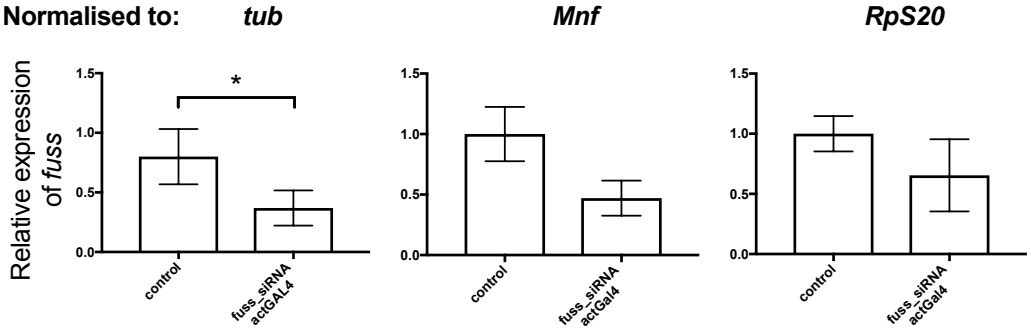
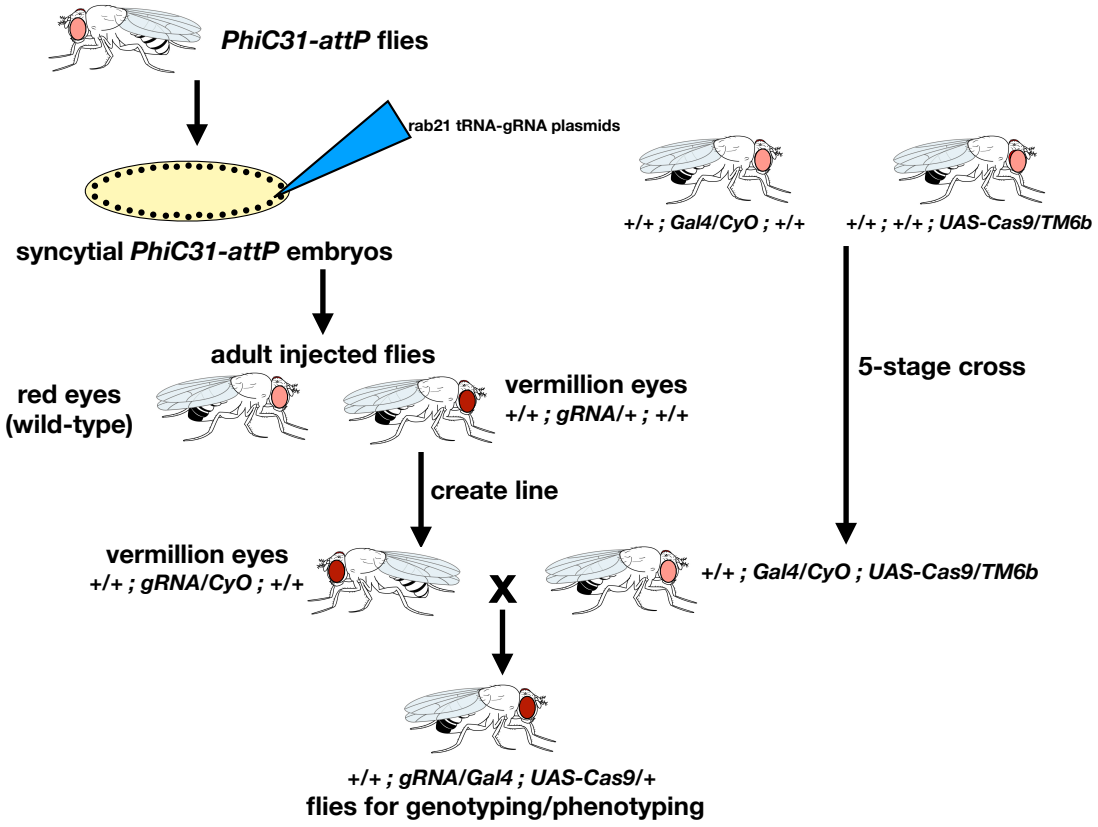


Fig. 5.14 A decrease in expression of *fuss* was detected after ubiquitous RNAi. Mean ± SEM is plotted of 5 repeats each containing 10 flies. Results were compared by Student's t-test. * p<0.05

5.8 GAL4-mediated CRISPR

One possible explanation for the lack of RNA perturbation seen with tissue-specific CRISPR is that the Cas9 was not expressed at high enough levels to cause mutation of the target genes. GAL4 can drive very high levels of expression of Cas9, which may increase the likelihood of mutation of target genes (Port et al., 2014; Port and Bullock, 2016). In addition, an elav-GAL4 line was already available (as used in chapters 3 and 4) for expression of GAL4 in neurons.

Therefore a strategy in which GAL4 was used to drive tissue-specific CRISPR was explored. This strategy is summarised in the following figure:



A line was obtained from the University of Cambridge fly facility which expresses Cas9 under the control of a UAS_G promoter (+/+ *;* +/+ *;* *UAS_G-Cas9/TM6b*).

Lines were also used which express GAL4 neuronally (+/+ *; elav-Gal4/CyO ;* +/+) and throughout the whole body (+/+ *; actin5C-Gal4/CyO ;* +/+).

Using a series of crosses, two sets of flies were created which express both GAL4 and UAS_G-Cas9, meaning that the expression pattern of Cas9 is controlled by that of the GAL4 promoter:
+ / + *; elav-GAL4 / CyO ;* *UAS_G-Cas9 / TM6b*
and
+ / + *; actin5C-GAL4 / CyO ;* *UAS_G-Cas9 / TM6b*

The intention was to establish stable lines of these two genotypes and cross them to the gRNA flies created in section 5.7 which ubiquitously express the gRNAs. However, the GAL4;Cas9 flies had very low viability and were infertile, preventing any further use of this method. By comparison, the parental lines were all viable and fertile.

5.9 Discussion

One of the main advantages of using *Drosophila* as a model organism is the ready availability of many different types of stocks for genetic manipulation of the genes of interest. However, these mass-produced lines do have some problems with off-target effects. Therefore in this chapter, the practicality of using CRISPR to generate specific *de novo* mutations was explored to use for confirmation of results from the RNAi screen.

5.9.1 Homogeneous CRISPR

In the first method tested (section 5.6) , an approach involving both tRNA-mediated liberation of gRNAs and co-mutation of *white* was used to facilitate mutation of the target genes. To this end, plasmids encoding multiple gRNAs to target *rab21*, *fuss* and *white* were successfully created. Injection of these plasmids into Cas9-expressing embryos led to creation and identification of three lines with mutation of *rab21* (and *white*), but none with mutation of *fuss*.

These methods were successful in that mutants were successfully created. However, since the flies which retained red eyes were not genotyped, it is not known what proportion (if any) had mutations in *rab21* and therefore it is impossible to tell whether the co-injection strategy facilitated the screening process as has been published (Ge et al., 2016). Similarly, whether encoding multiple gRNAs on the same plasmids by separating them with tRNAs was more efficient than making separate plasmids for each gRNA is unknown.

The intention was to delete all of the DNA in between the two gRNA sites. However, this was not detected in any of the flies screened. Instead, for *rab21* all of the mutations found were at just one gRNA site. Luckily, the functional gRNA was the one closer to the start of the gene, increasing the chance that the mutations would perturb the gene. This shows that the CRISPR strategy worked but suggests that there was a problem with the gRNAs.

It has been shown that the position of the gRNA on the plasmid does not affect the mutagenesis efficiency (Port and Bullock, 2016). Instead, the most likely explanation for the failure to mutate the second site is that the gRNA design was not optimal. In the future, this may not be a problem as the algorithms to design gRNAs improve. In addition, more than one gRNA should be used at both target sites. It is possible to use the same tRNA approach to create a plasmid containing up to four gRNAs, although the cloning requires more stages and so is less time-efficient.

Mutation of *rab21*

Two *rab21* mutant lines were created. In the $rab21^{\Delta 13}$ line, a premature stop codon results in deletion of most of the protein including the GTPase domain, and so would be predicted to strongly affect gene function. Indeed, qPCR showed that RNA levels of *rab21* are decreased (possibly due to nonsense-mediated decay of the mRNA transcripts) and a phenotype was seen in many of the energy homeostasis assays. In the $rab21^{\Delta 6}$ line, most of the protein is still present but it is possible that the mutation would influence folding or regulation of the *rab21* mRNA/protein. Therefore, any effects would be expected to be milder for the $rab21^{\Delta 6}$ than for the $rab21^{\Delta 13}$ line, and this is seen in both qPCR and some of the phenotype assays. There is no premature stop codon in the $rab21^{\Delta 6}$ line and so reasons for the decrease in mRNA levels are not obvious, but it could be hypothesised that there is some form of feedback loop.

The RNA levels are affected less by the CRISPR mutations than they are by RNAi. This is unsurprising given that RNAi targets the RNA and qPCR measures RNA levels, whereas CRISPR affects the DNA and protein. However, it does not follow that the effects on phenotype will be larger for RNAi than for CRISPR. If the gRNAs had been designed to target the promoter or other regulatory sequences of the *rab21* gene, then a larger change in RNA levels would be expected with CRISPR than was observed.

The phenotype of the homogeneous CRISPR flies matched that of the ubiquitous RNAi flies for *rab21*. Since it is unlikely that the two models have the same off-target effects, this provides validation that both genetic approaches are successfully targeting *rab21*. Specifically, the dye ingestion, mass and CAFE results matched precisely. Results for which there was a trend with RNAi (climbing and starvation resistance) became significant with CRISPR, either because the genetic perturbation is stronger and/or because the CRISPR flies were all genotyped within the same batch which decreased variation in the results compared to the RNAi flies which were phenotyped across multiple batches. The only disagreement in the results is the glucose assay where no phenotype is seen for CRISPR but a decrease is seen for RNAi. It is possible that this last result is a false positive in the RNAi flies due to to off-target effects of the siRNA, or a false-negative result in the CRISPR flies due to a problem with the assay.

Failure to mutate *fuss*

As stated in section 5.4, one of the reasons that *fuss* was chosen as a target gene is that its presence on chromosome 4 makes it difficult to target (Sun et al., 2000) and so the failure to create a mutant line with homogeneous CRISPR is not surprising.

Only a very small proportion (9%) of the injected embryos survived to adulthood which reduced the chance of finding mutants. Since the Cas9 expression was limited to the pole cells, no (or very few) mutations should have occurred in the somatic cells and so mutation of *fuss* is unlikely to be the cause of the lethality. This suggests that instead there was a problem with the microinjection, either molecularly (for example the injected plasmid solution was too concentrated) or physically (for example a blunt needle). For the few injected embryos which did survive to adulthood, there were no white-eyed offspring.

Further, the chromosome 4 balancer lines were very unhealthy. Therefore, even if a mutation in *fuss* had been found it may have been difficult to maintain the line if the mutation affected viability or fertility. As chromosome 4 contains less than 0.5% of a fly's genes (Sun et al., 2000; Van Holde and Zlatanova, 2018), the need for chromosome 4 balancers in future work is likely to be low, but if required, the maintenance conditions would need to be optimised.

5.9.2 Failure of tissue-specific CRISPR

Perturbation of neither *fuss* nor *rab21* could be shown using tissue-specific CRISPR (section 5.7). This is in agreement with published results which found that CRISPR using tissue-specific expression of Cas9 combined with ubiquitously expressed sgRNAs often leads to poorly penetrant phenotypes (Port et al., 2014; Port and Bullock, 2016).

For *rab21*, the problem is not the gRNAs because the same gRNAs were able to cause mutation with the homogeneous CRISPR method. Further, the gRNA plasmid is definitely being expressed within the genome as evidenced by the presence of vermillion eye colour. In addition, the qPCR protocol is functional as it was able to detect differences in RNA levels for both RNAi and homogeneous CRISPR. Therefore, the exact reasons why no perturbation was seen remain unclear. It could be speculated that CRISPR is less efficient in adults than in embryos, perhaps because the fast nuclear division occurring in embryos provides more opportunities for Cas9 to access the target DNA. Alternatively, perhaps the Cas9 used for tissue-specific CRISPR is less active than that which was used for homogeneous CRISPR. In addition, the background controls used in the qPCR were the two parental lines — these are

homozygous throughout their genome unlike their heterozygous offspring so are perhaps not ideal for use as background controls.

When the fuss-gRNA line was crossed to act-Cas9, qPCR showed an increase in the level of *fuss* mRNA. Again, the reasons for this unexpected result are not certain, but in addition to the theories discussed above for *rab21*, it could be hypothesised that because *fuss* is on heterchromatic chromosome 4, the gRNA-directed Cas9 may cause the chromatin structure around *fuss* to relax and thus increase gene transcription.

For both genes, it is possible that mutations were in fact being created, but since this could not be shown it would be difficult to confidently interpret any phenotype data.

5.9.3 Failure of GAL4-mediated CRISPR

The intention was to cross GAL4;UAS_G-Cas9 lines to gRNA-expressing flies in order to use the high expression levels of the UAS_G-GAL4 system to increase the likelihood of obtaining CRISPR mutations (section 5.8). However these lines had very low viability and were infertile. The four starting parental lines were healthy, meaning that it is not the individual transgenes that are problematic. In theory, their combination should also not cause viability or fertility issues since no CRISPR-induced mutation should occur in the absence of gRNAs. This suggests that the expression of this Cas9 is having a detrimental effect even in the absence of gRNAs. Published results have also found that GAL4-mediated expression of UAS_G-Cas9 causes lethality (Port et al., 2014; Port and Bullock, 2016; King-Jones and Huynh, 2018) and the authors' suggestion was that the very high levels of Cas9 expression that result from the UAS_G-GAL4 system causes substantial cytotoxicity independent of the endonuclease activity.

5.9.4 Comparison of homogeneous and tissue-specific CRISPR

Tissue-specific CRISPR has a number of advantages in comparison to homogeneous CRISPR, and many of these are the same as the differences between LoF and RNAi discussed in chapter 4. Firstly, the information is tissue-specific which both provides mechanistic information about the gene being studied and avoids any detrimental effects of a whole-organism LoF. Secondly, the gRNA and Cas9 lines are maintained separately which allows mutations in a range of different tissues or conditions to be studied without needing to create an additional line. Finally, the post-injection screening is easier for tissue-specific CRISPR as it requires

only looking at eye colour for plasmid integration rather than the sequencing needed to detect mutation by homogeneous CRISPR.

There are also four disadvantages of the tissue-specific method compared to homogeneous CRISPR. Firstly, CRISPR almost always has off-target effects due to sequence similarity with non-target genes (Zhang et al., 2015b) and unlike with homogeneous CRISPR, these off-targets cannot be removed by crossing. Secondly, suitable lines to use as background controls are not easily available for tissue-specific CRISPR but are for homogeneous CRISPR. Further, tissue-specific CRISPR may create different mutations in every cell which makes the results more difficult to interpret than homogeneous CRISPR where the exact mutation is known. For homogeneous CRISPR, different lines can have different mutations which potentially provides different information. Finally and most importantly, homogeneous CRISPR was demonstrated to cause perturbation of the gene of interest whereas this could not be shown for tissue-specific CRISPR.

5.9.5 Comparison of RNAi and CRISPR

The main draw-back of using RNAi is the potential for off-target effects. Like RNAi, CRISPR can also have off-target effects. However, it is unlikely that the off-targets would be the same for both CRISPR and RNAi, and so the results can be compared to look for specific and non-specific phenotypes.

Creation and screening of the CRISPR lines is very low-throughput compared to RNAi. siRNA lines are already available and can be shipped and passed through quarantine in four weeks. By comparison, it took one year from selection of the target genes to obtaining homogeneous CRISPR flies for phenotyping. In this time, a mouse model of the gene of interest could be obtained, which may provide information that is more directly relevant to human physiology. The tissue-specific CRISPR gRNA flies were obtained more quickly (approximately 1 month) after injection so could provide results in a time-frame more appropriate for a high-throughput screen.

Two genes were chosen for CRISPR targeting: *rab21* and *fuss*. The homogeneous CRISPR only worked for *rab21* and the tissue-specific CRISPR did not appear to work for either. Although most lines studied in this thesis were not analysed by qPCR to check for knock-down of gene expression, RNAi is widely accepted within the *Drosophila* research community as a reliable tool (for example see St Johnston (2002); Mohr et al. (2014); Wangler et al. (2017) and Heigwer et al. (2018)). Since the idea of a *Drosophila*-based screen is to simultaneously study a very large number of genes to select top candidates, a few non-

functional siRNA lines (and therefore false negative results) does not matter. In comparison, the low-throughput nature of CRISPR means that it is only suitable for studying a small number of genes, and so it is important that the genetic perturbation is strong, accurate, reliable and precise, which it was not possible to show here.

The aim of this thesis was to generate a model for high-throughput and economical analysis of large numbers genes relevant to human obesity. Based on the work in this and previous chapters, tissue-specific RNAi in *Drosophila* seems to be the best model for such a purpose.

On the other hand CRISPR has potential for studying a small number of genes in detail and so could be used to study just the top hits of an RNAi-based screen. In that case, further optimisation of the methods would be required including: use of more than two gRNAs to target each gene; creation of lines more suitable to use as background controls for neuron-specific CRISPR; and acquisition/creation of neuron-specific Cas9 lines. It would perhaps be helpful to use mutation of a gene such as *white* which has an easily scored phenotype to carry out the optimisation, before moving to genes like *fuss* and *rab21* which require molecular techniques for assessment.

5.10 Summary

In summary, in this chapter CRISPR was explored as a method of perturbing expression of genes of interest. Use of tRNAs to separate gRNAs was used to facilitate creation of a single gRNA-containing plasmid for each target gene. Co-targeting of the *white* gene was used to facilitate the process of screening for mutants. Two mutant lines were successfully created for *rab21* and shown to have decreased levels of *rab21* mRNA. Phenotypes of these lines matched those of flies with ubiquitous RNAi against *rab21*. No mutant lines were successfully created for *fuss*. Further, attempts to do tissue-specific CRISPR were thwarted by the inability to detect whether mutation was occurring or not. Thus for now, neuronal RNAi remains the most suitable model for the screen to probe genes of interest for a role in the CNS's control of food intake and energy homeostasis.

Chapter 6

A *Drosophila*-based screen of hypothalamic genes affected by fasting

As discussed in chapter 1, up to 70% of the variation in BMI and fat mass between individuals is influenced by genetic factors (Maes et al., 1997). However to date, genetic studies such as GWAS have only been able to account for less than 10% of this variation (Bogardus, 2009; Speakman et al., 2018).

Many of the genes that have been studied from Mendelian obesity and polygenic common obesity are linked to the leptin-melanocortin pathway in the CNS which has a key role in the control of energy homeostasis (see chapter 1). Therefore a different approach to studying obesity genetics is to look specifically at the POMC and NPY neurons in the arcuate nucleus, since these are an important site of action of the leptin-melanocortin pathway. Briefly, in response to energy deficit, NPY neurons cause an increase in food intake brought about by changes in gene expression and electrical activity, whereas POMC neurons have the opposite action during energy surplus.

Therefore in this chapter, a data set derived from transcriptomic data of mouse POMC and NPY neurons subjected to metabolic perturbation (fasting) was interrogated using the high-throughput *Drosophila*-based screen for energy homeostasis phenotypes which was developed and refined in chapters 3-5.

6.1 Transcriptomic studies of hypothalamic neurons

Gene expression is not uniform but instead has strong temporal and spatial characteristics. These depend on genetics and also on the present and past environment. For example, food intake affects the neuronal pathways that control feeding, but this would not be detectable by purely genomic data. The transcriptome refers to the RNA transcripts produced from the genome in a specific condition and/or cell. Transcriptomic techniques such as microarrays and RNAseq measure levels of these RNAs and their temporo-spatial variation. Transcriptomic studies generally produce large lists of data and so require a high-throughput model to study the identified genes.

Fasting increases hunger and triggers feeding behaviour, and so studying changes in the transcriptome bought about by fasting provides a foundation for further investigation into the regulation of energy homeostasis.

Henry et al. (2015) studied gene expression in both NPY and POMC neurons from fed and fasted mice. Briefly, they bred Npy^{hrGFP} and $Pomc^{topazFP}$ transgenic male mice from which the NPY and POMC neurons, respectively, could be isolated by their expression of fluorescent protein. For each genotype 5 young adult male mice (6.5-8 weeks) were fed *ad libitum* for 24 hours and 5 were deprived of food for 24 hours. Then the hypothalamus was dissected from each brain, the neurons dissociated, and Fluorescence Activated Cell Sorting (FACS) used to separately collect NPY and POMC neurons. From each of the 20 neuron pools, each containing 44-214 neurons, the RNA was extracted and RNAseq used to measure the expression levels of every gene. Henry et al. showed their samples were uncontaminated by checking for marker genes of other cells types, and that their results were accurate by checking that the data was consistent with previously published data for specific genes.

6.2 Chapter aims

The aim of this chapter was to use the *Drosophila*-based screen that was developed and validated in chapter 3 and tested in chapter 4 to probe the transcriptomics data from Henry et al. (2015) in order to identify top candidate genes that warrant further exploration in mammalian models.

6.3 Chapter-specific methods

In previous chapters, the flies were housed in incubators with 60-70% humidity. For the work done in this chapter, the humidity regulator was broken and so the flies were instead kept at 10-40% humidity. This dried the food, making it more difficult for larvae to infiltrate and feed which, in turn, may affect their development and physiology. To compensate, 10 UAS_G-siRNA virgin female flies were used for each cross (rather than 5).

Since all of the genes in this data set are known to be expressed in the mammalian brain, all were phenotyped in *Drosophila* using neuronal RNAi. Only viability was tested using ubiquitous RNAi.

The climbing and starvation assays were not used as they were found in chapters 3 and 4 to be less sensitive than the other screen assays. Instead the TAG levels were determined as an extra measure of physiology, but using the method from Baranski et al. (2018). Their method involves no centrifugation of the fly homogenate sample, unlike the method detailed in chapter 2 which involves two centrifugation steps. As can be seen in figure 6.1, Baranski et al.'s method resulted in much higher levels of TAG in each sample.

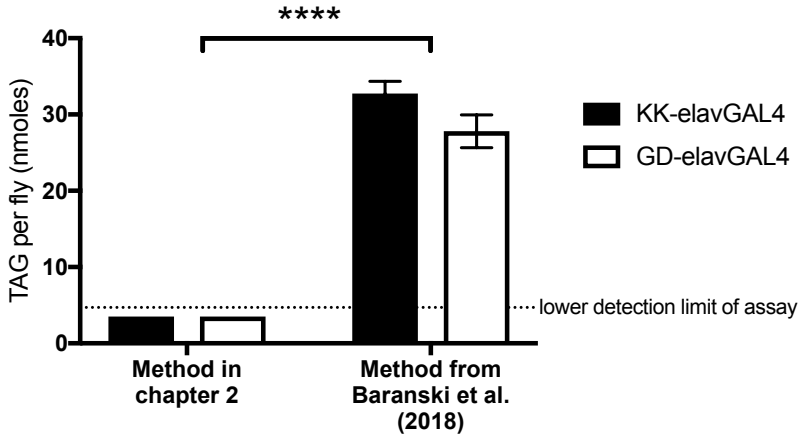


Fig. 6.1 TAG levels in KK/GD-elavGAL4 samples are much higher using the method described by Baranski et al. (2018) than they are from the method described in chapter 2 of this thesis. Mean ± SEM is plotted of 10 repeats each containing 15 flies. Groups were compared by two-way ANOVA. **** $p<0.0001$

The scoring formula was amended to accommodate these changes in the screen assays:
$Score = 1 - (\frac{1}{15})(5p_{CAFE} + 4p_{dye} + 3p_{TAG} + 2p_{mass} + p_{glucose})$

With the initial set of assays that were used in chapters 3 and 4, the positive and negative control genes could be differentiated by having a score above or below 0.7, respectively. Since the formula above uses a slightly different suite of assays, it is possible that the threshold score for a gene to be considered a hit should be amended. However, as an indicator that 0.7 is still appropriate to use, if the scores for the positive and negative control genes are calculated with the CAFE, dye food over-feeding, wet mass and glucose assay results (i.e. the same as this new formula but without the TAG), then 0.7 remains the distinction between the groups of controls.

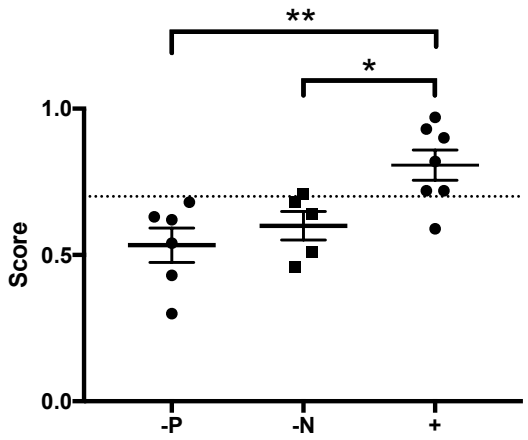


Fig. 6.2 The positive control genes are distinguishable from the negative control genes with the new scoring system. -P are peripheral disease negative control genes, -N are neuronal disease negative control genes, and + are positive control genes (see chapter 3). Groups were compared by Student's t-test. * $p<0.01$, ** $p<0.001$.

Results

6.4 Selection of genes for phenotyping

In total, Henry et al. reported the expression levels of 35266 mouse genes in both POMC and NPY neurons. The process used to select genes from this list for screening in *Drosophila* is detailed below and summarised in figure 6.3.

Of all the mouse genes, the number which showed a significant change ($p<0.05$) in expression between the fed and fasted states was 1234 in POMC cells and 3655 in NPY cells.

To further filter the data, only those genes which show at least a 1.5-fold change in expression were studied because this increases the statistical power of the data (Conesa et al., 2016). Most of the genes filled this criterion, resulting in removal of only 196 genes from the POMC cell list and 101 from the NPY cell list.

Since POMC and NPY neurons have opposing roles in the control of energy homeostasis (see chapter 1), any genes which change in a reciprocal fashion in these neurons are likely to be the best targets for future therapeutic development. Fasting caused expression of 39 genes to decrease in POMC cells and increase in NPY cells. Conversely, fasting caused expression of 153 genes to increase in POMC cells and decrease in NPY neurons.

All of these mouse genes had a strong homologue in humans which is important if the results are to be used to understand human physiology.

Next, ENSEMBL was used to identify *Drosophila* orthologues of these mouse genes (Zerbino et al., 2018). 58% had a *Drosophila* orthologue, which is approximately the same as the genome-wide average of 53% (Wangler et al., 2017). Most of the mouse genes had just one homologue but some had multiple, and in one case two mouse genes had the same fly homologue. Overall, this gave a list of 157 *Drosophila* genes.

To increase the likelihood of these *Drosophila* genes being functional homologues of the mouse genes, they were only taken further if the sequence homology was greater than 30%. This brought the total to 91 genes. Further the genes had to be expressed in the adult and/or larval CNS according to published microarray data (Graveley et al., 2011), bringing the total to 61 genes. The neuronal expression of the genes also ensured that it was appropriate to study the genes using neuron-specific RNAi.

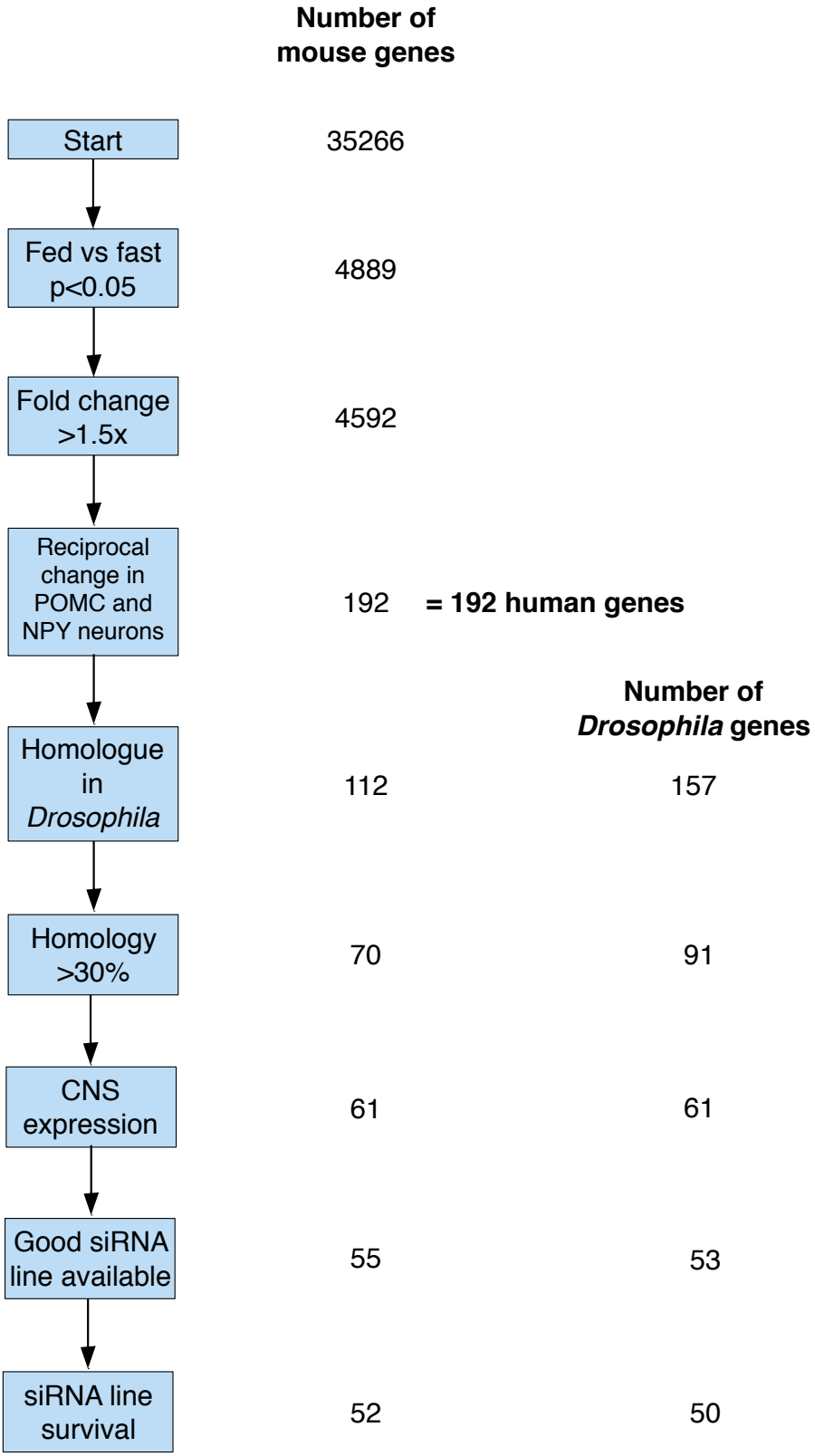


Fig. 6.3 Criteria used to select genes for study in *Drosophila* from genetic data by Henry et al. (2015)

Finally, UAS_G-siRNA lines to target the *Drosophila* genes were ordered from VDRC. For 2 of the genes, no siRNA line was available. For 6 of the genes, the only lines available had multiple off-target effects and so these genes were not ordered. As before, GD lines were ordered in preference to KK lines. In total 53 *Drosophila* siRNA lines were ordered.

12 siRNA lines were ordered representing 13 mouse genes whose expression goes up in NPY neurons and down in POMC neurons during fasting. Given the role that NPY neurons play in triggering feeding, and that for most genes an increase in expression is more important for their function than a decrease, these 13 genes are referred to in this thesis as orexigenic genes.

41 siRNA lines were ordered representing 40 mouse genes whose expression increases in POMC neurons and decreases in NPY neurons during fasting. These 40 genes are referred to in this thesis as anorexigenic genes.

3 of the siRNA lines ordered died of a mite infection in quarantine before they could be phenotyped.

Therefore, in total, 50 genes were studied, which are listed in table 6.1.

Table 6.1 Mouse genes which respond to fasting in a reciprocal manner in POMC and NPY neurons, and their *Drosophila* orthologues. [m] Mouse genes with the same *Drosophila* orthologue. [f] Fly genes with the same mouse orthologue.

RNAseq of mouse hypothalamic neurons (Henry et al., 2015)							*Drosophila* orthologue	
Gene	POMC neurons			NPY neurons				
	Fast - Fed (TPM)	t test p value	Fold change	Fast - Fed (TPM)	t test p value	fold change	Gene	Homology (%)
Orexigenic genes								
Arpc5l	-62.6	0.0210	1.5	54.5	0.0130	1.5	*Arpc5*	44
Hsd17b12	-35.9	0.0021	2.0	46.3	0.0000	2.1	*spidey*	44
Tubb4b	-97.7	0.0222	1.6	112.7	0.0010	1.6	*βTub56D*	94
Chml	-1.8	0.0407	2.6	3.6	0.0247	1.6	*Rep*	33
Ddah1	-94.4	0.0069	1.8	79.1	0.0014	2.6	*CG1764*	38
Fam162a	-42.9	0.0159	1.9	65.0	0.0348	2.0	*CG9231*	32
Gem	-19.7	0.0357	2.3	74.9	0.0004	4.9	*Rgk1*	30
Hcrtr2	-15.3	0.0054	3.1	11.7	0.0046	6.6	*SIFaR*	30
Prkcsh	-30.0	0.0481	1.5	49.6	0.0447	1.7	*GCS2β*	38
Spred2 [m]	-40.5	0.0327	2.4	69.3	0.0017	1.9	*Spred*	35
Spred3 [m]	-46.1	0.0029	6.1	70.0	0.0000	2.7	*Spred*	30
Ssr4	-122.4	0.0158	2.0	101.1	0.0127	1.6	*Tapδ*	34
Anorexigenic genes								
Acvr2a	58.5	0.0099	1.7	-31.9	0.0088	1.5	*put*	47
Aldh6a1	35	0.0044	2.2	-15.3	0.0373	1.5	*CG17896*	67
Atp2b2	42.7	0.0001	1.5	-50.5	0.0007	1.7	*PMCA*	60
Cacnb4	12.2	0.0002	1.5	-36.3	0.0001	2.3	*Ca-β*	53
Cdk14	30.7	0.0160	1.6	-12.8	0.0025	4.6	*Eip63E*	52
Dtna	105.9	0.0004	1.9	-133.2	0.0000	2.2	*Dyb*	43
Elavl2 [f]	14.0	0.0061	1.5	-32.8	0.0016	2.0	*Rbp9*	58
Elavl2 [f]	14.0	0.0061	1.5	-32.8	0.0016	2.0	*fne*	59
Epb4.1l4b	13.1	0.0050	1.9	-22.1	0.0001	4.2	*yrt*	47
Gda	38.2	0.0045	1.7	-21.2	0.0352	2.1	*DhpD*	45
Gpn3	51.6	0.0009	2.9	-24.1	0.0062	1.8	*CG2656*	51
Nalcn	58.6	0.0016	1.7	-49.5	0.0015	1.6	*na*	57
Nedd4l	33.6	0.0102	1.7	-35.5	0.0003	1.7	*Nedd4*	49
Nmt1	20.8	0.0057	1.5	-21.9	0.0445	1.5	*Nmt*	56
Pofut1	7.6	0.0014	2.4	-6.2	0.0295	3.4	*O-fut1*	42
Prkaa1	23.4	0.0284	1.5	-21.3	0.0190	1.5	*AMPKα*	60
Pura	83.5	0.0206	1.5	-111.8	0.0137	1.9	*Pur-α*	44
Pygl	11.5	0.0285	1.8	-15.1	0.0003	2.8	*GlyP*	71
Rasgef1a [f m]	29.8	0.0055	1.5	-15.6	0.0312	1.5	*CG7369*	41
Rasgef1c [f m]	5.4	0.0473	1.9	-9.6	0.0082	9.5	*CG7369*	46
Rasgef1a [f m]	29.8	0.0055	1.5	-15.6	0.0312	1.5	*CG4853*	36
Rasgef1c [f m]	5.4	0.0473	1.9	-9.6	0.0082	9.5	*CG4853*	37
Rps6ka6	12.7	0.0277	2.1	-40.1	0.0138	2.4	*S6kII*	49
Rufy2	34.1	0.0172	1.5	-39.3	0.0281	1.7	*CG31064*	42
Syt7	4.3	0.0361	2.9	-13.9	0.0085	1.9	*Syt7*	46
Tub	97.9	0.0083	1.6	-125.9	0.0002	2.2	*ktub*	42
Ccdc132	20.5	0.0471	1.6	-16.8	0.0101	1.5	*Vps50*	31
Fat3	15.7	0.0163	2.0	-28.3	0.0019	6.7	*kug*	36
Htr1a [f]	12.1	0.0346	5.4	-30.9	0.0000	4.9	*5-HT1A*	36
Htr1a [f]	12.1	0.0346	5.4	-30.9	0.0000	4.9	*5-HT1B*	35
Kcnj3	37.5	0.0431	1.5	-18.6	0.0338	2.1	*Irk2*	35
Kcnq3	90.7	0.0066	3.1	-54.3	0.0261	2.5	*KCNQ*	31
Lipa	26.0	0.0061	1.6	-19.9	0.0002	2.4	*CG18301*	34
Lrch2	36.5	0.0009	1.5	-48.6	0.0259	1.8	*Lrch*	31
Nr1d2	67.2	0.0024	2.8	-66.8	0.0003	2.3	*Eip75B*	30
Rbfox1	40.1	0.0046	1.6	-23.1	0.0004	2.5	*Rbfox1*	36
Rps6ka5	22.1	0.0008	2.7	-31.3	0.0003	2.2	*JIL-1*	38
Slc44a5	15.4	0.0256	3.0	-6.0	0.0328	1.7	*Ctl2*	34
Smpdl3b	4.4	0.0415	3.9	-5.1	0.0080	5.9	*CG32052*	32
Sytl4	11.0	0.0330	2.4	-6.9	0.0004	3.0	*btsz*	35
Vps39	46.6	0.0072	1.6	-56.1	0.0044	1.7	*Vps39*	36

6.5 Rapidity of the *Drosophila*-based screen

The primary aim of this thesis was to develop a high-throughput *in vivo* screen of feeding behaviour and energy homeostasis phenotypes in *Drosophila*. The study of the fasting transcriptomics genes in this chapter allowed the speed of the screen to be accurately measured. From a start-point of being able to collect siRNA virgin flies, it took 6 weeks to phenotype 50 lines (plus 3 background controls) using 5 assays each.

6.6 Viability of the transcriptomic genes in *Drosophila*

The lethality of each gene with whole body RNAi can be seen in figure 6.10. Of the GD lines, 31% are lethal with ubiquitous RNAi. This includes one line in which the flies eclose but die within a few days. This lethality is only marginally higher than the genome-wide average of 25% (Wangler et al., 2017).

Of the KK lines, 7 are viable, 1 is viable only in females, and 9 are lethal with ubiquitous RNAi, although some of this lethality may be caused by over-expression of *tiptop* (see section 6.7).

6.7 Lines affected by *tiptop* over-expression

Of the 16 siRNA lines with a KK background, 7 were lethal during / before the pupal stage when crossed to act-GAL4. Data from these genes was therefore compared to both KK-elavGAL4 and KK_{tiptop}-elavGAL4 as a background control. In each case the smaller phenotype was chosen as the result in order to avoid false positive results which are more detrimental than false negative results for the purpose of candidate gene selection. The results are shown in figure 6.4.

The siRNA lines which target *betaTub56D, Rbfox1, 5-HT1A, 5-HT1B* and *fne* have a smaller phenotype using a KK background, with both fewer significant assay results and a lower overall score.

Only one line (*Eip75B*) has a smaller phenotype with KK_{tiptop} and was thus assigned to this background.

For *DhpD*, there are significant results in more assays when compared to KK_{tiptop}, but the score is lower with KK, and so the appropriate background is unclear. For data presented later in this chapter the KK background was used.

Drosophila Gene	Human Gene	siRNA line	Background	CAFE	Dye Food	TAG	Wet Mass	Glucose	Score
betaTub56D	Tubb4b	VT1	KK	0.5987	0.0869	0.7759	0.0707	**0.0051**	0.61
			KK_{tiptop}	**0.0210**	0.3917	**0.0000**	0.4757	**0.0034**	0.82
Rbfox1	Rbfox1	VT48	KK	0.8994	0.0502	0.0866	0.7064	0.0360	0.57
			KK_{tiptop}	**0.0024**	0.3917	**0.0000**	0.6009	0.6527	0.77
5-HT1A	Htr1a	VT51	KK	0.5936	0.2340	**0.0096**	0.3103	**0.0117**	0.70
			KK_{tiptop}	**0.0016**	0.2426	**0.0000**	0.2931	0.6441	0.86
5-HT1B	Htr1a	VT52	KK	0.0773	0.5637	0.0917	**0.0320**	0.8682	0.74
			KK_{tiptop}	**0.0437**	0.0750	**0.0001**	0.1306	**0.0005**	0.86
fne	Elavl2	VT54	KK	0.3685	0.5271	0.0633	0.0869	**0.0001**	0.71
			KK_{tiptop}	**0.0165**	0.0609	**0.0000**	0.9255	**0.0160**	0.86
DhpD	Gda	VT39	KK	0.6468	**0.0005**	**0.0151**	**0.0316**	**0.0135**	0.78
			KK_{tiptop}	**0.0019**	0.3349	**0.0000**	0.4041	**0.0035**	0.86
Eip75B	Nr1d2	VT44	KK	**0.0252**	**0.0241**	0.0025	**0.0161**	**0.3633**	0.96
			KK_{tiptop}	0.1337	0.5431	0.0007	**0.0125**	**0.0001**	0.81

Fig. 6.4 p values of potential *tiptop*-affected lines compared to both KK-elavGAL4 and KK_{tiptop}-elavGAL4. The background with the smaller phenotype was chosen, indicated in orange. Red is an increase compared to controls, blue is a decrease, and significant results are highlighted

6.8 Phenotype of transcriptomic genes

The feeding and physiology phenotypes of flies with neuronal RNAi of each of the mouse hypothalamic transcriptomics genes are shown in figures 6.5 to 6.9, with the overall scores shown in figure 6.10. As can be seen, each assay detected multiple significant changes, and many nearly significant changes. Almost all of the significant increase phenotypes are seen in the anorexigenic genes (genes whose expression decreases in NPY neurons during fasting), rather than in the orexigenic genes (whose expression decreases in POMC neurons during fasting). The scores have a large range (0.34 to 0.98) meaning that the top hit genes can clearly be distinguished from those at the bottom.

Genes which were not hits

Based on their score (<0.7) and the individual assay results in the *Drosophila*-based screen, the following 31 mouse genes were not deemed worthy of further study: Epb4.1l4b, Gpn3, Rbfox1, Pygl, Ccdc132, Syt7, Lipa, Smpdl3b, Pura, Acvr2a, Vps39, Atp2b2, Kcnq3, Nalcn, Nmt1, Fat3, Prkaa1, Hcrtr2, Hsd17b12, Fam162a, Tubb4b, Ssr4, Arpc5l, Gem, Prkcsh, Spred2, Spred3, Ddah1, Chml, Rasgef1a and Rasgef1c.

From the genes with a score >0.7, there was one mouse gene (Nr1d2) which shows decrease phenotypes in assays of both energy intake and physiology. As has been mentioned in previous chapters, this can be a sign of unhealthiness rather than an energy homeostasis phenotype, and so this gene is likely not worthy of further study.

Genes with discordant results

In addition, there were 6 mouse genes with a score >0.7, but with discordant assay results in the *Drosophila*-based screen: Cacnb4, Dtna, Pofut1, and Rps6ka6 showed an increase in food intake but a decrease in measures of physiology, and Slc44a5 and Kcnj3 showed the opposite phenotype. These genes may be worthy of further study, but are not considered the top hits.

Top hit genes

Finally, there are 10 mouse genes for which neuronal RNAi in *Drosophila* causes an increase in both food intake and measures of physiology, giving an overall score >0.7. These genes are the top hits for further study: Lrch2, Tub, Cdk14, Aldh6a1, Rps6ka5, Nedd4l, Sylt4, Elavl2, Rufy2, and Htr1a.

CAFE assay

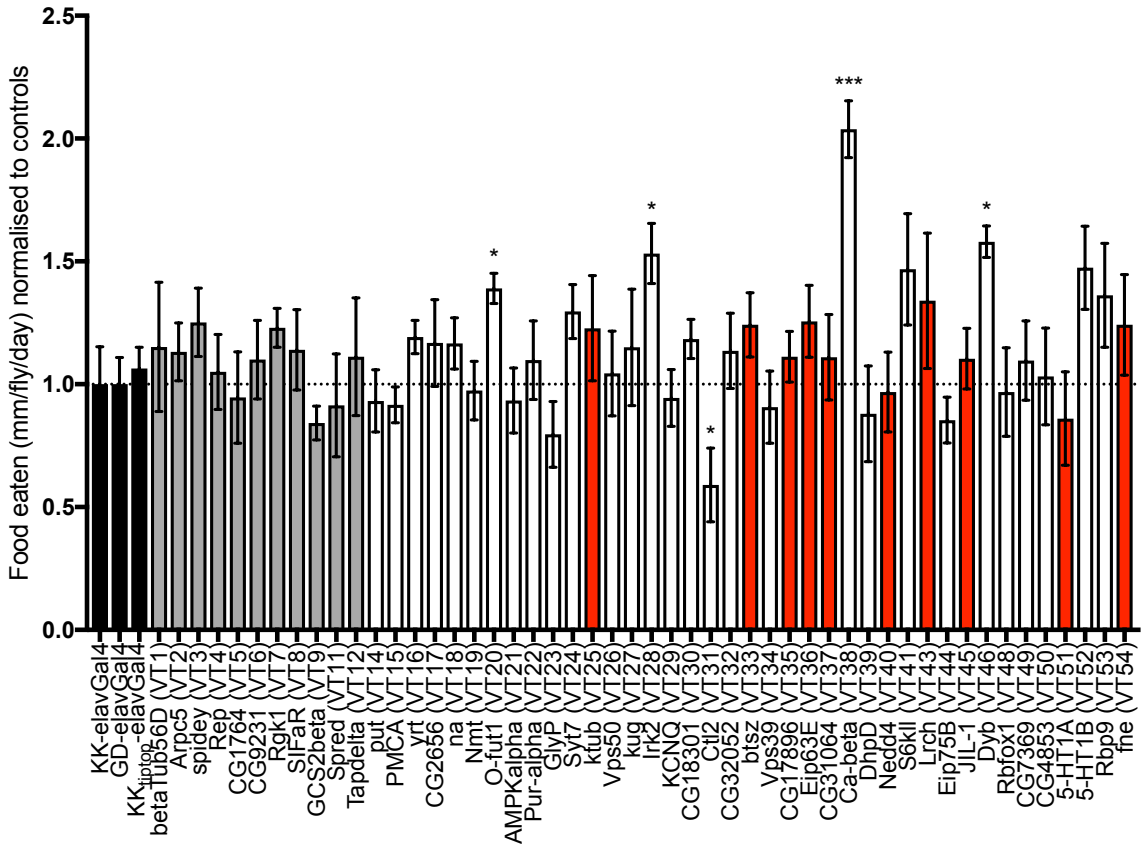


Fig. 6.5 Food intake with CAFE assay by flies with neuronal RNAi of transcriptomics-identified genes. Background control lines are black, orexigenic genes are grey, anorexigenic genes are white, and the 10 top hit genes are highlighted in red. Mean ± SEM is plotted of 5 repeats each containing 8 flies. Results were compared to appropriate background by Student's t-test. * $p<0.05$, ** $p<0.01$, *** $p<0.001$

Dye food over-feeding assay

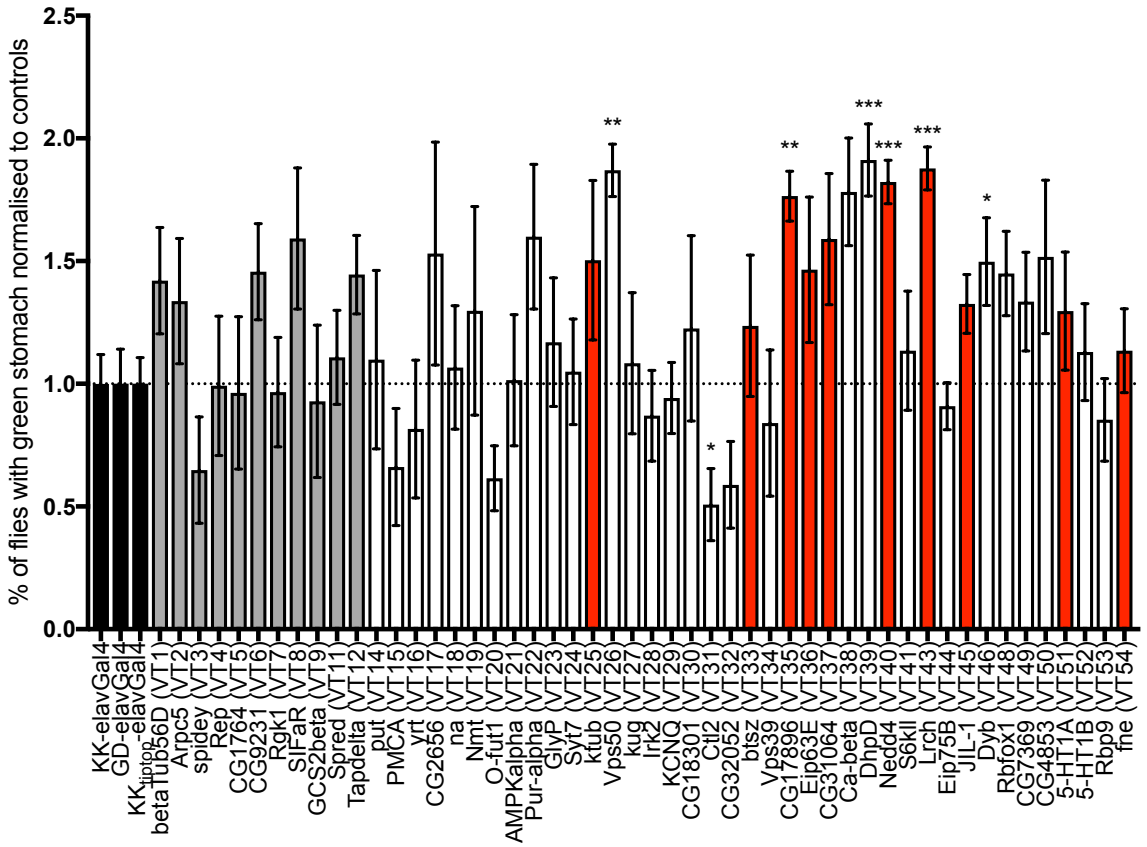


Fig. 6.6 Ingestion of dye food by flies with neuronal RNAi of transcriptomics-identified genes. Background control lines are black, orexigenic genes are grey, anorexigenic genes are white, and the 10 top hit genes are highlighted in red. Mean ± SEM is plotted of 5 repeats each containing 15 flies. Results were compared to appropriate background by Student's t-test. * p<0.05, ** p<0.01, *** p<0.001

TAG levels

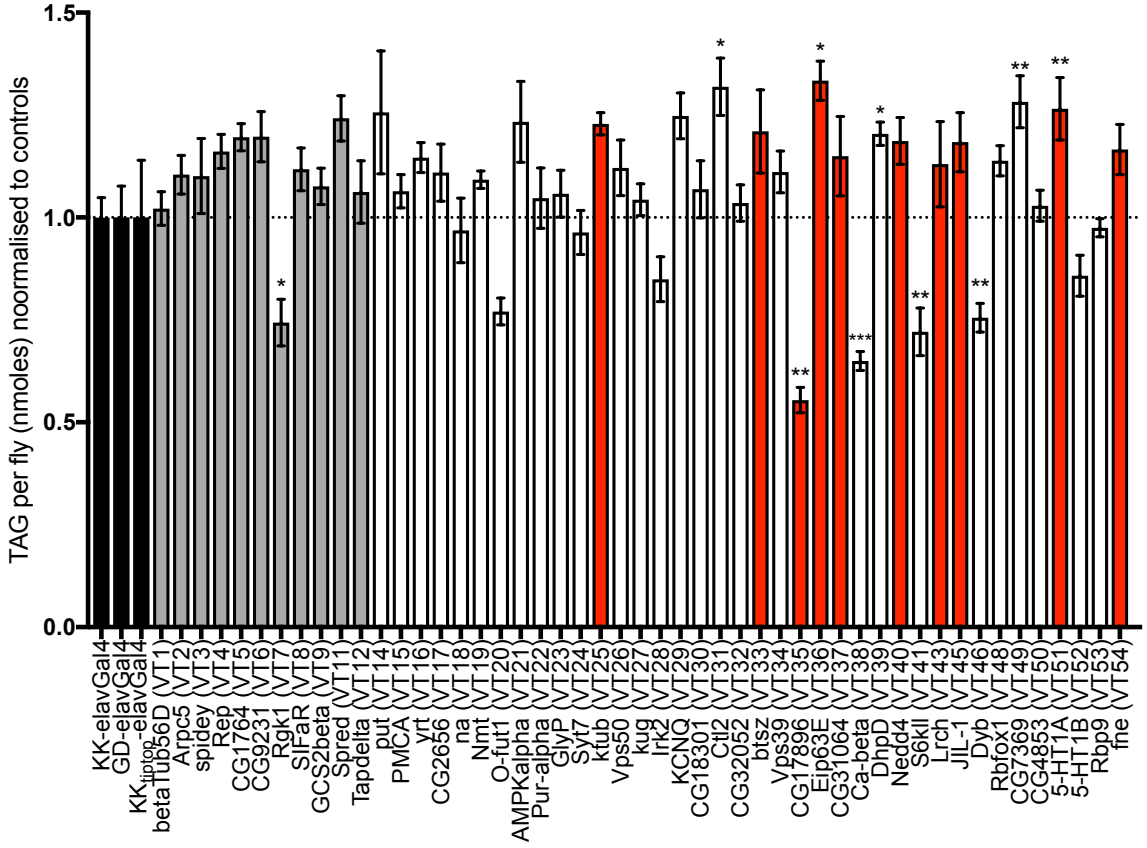


Fig. 6.7 TAG levels in flies with neuronal RNAi of transcriptomics-identified genes. Background control lines are black, orexigenic genes are grey, anorexigenic genes are white, and the 10 top hit genes are highlighted in red. Mean ± SEM is plotted of 5 repeats each containing 15 flies. Results were compared to appropriate background by Student's t-test. * p<0.05, ** p<0.01, *** p<0.001

Wet mass

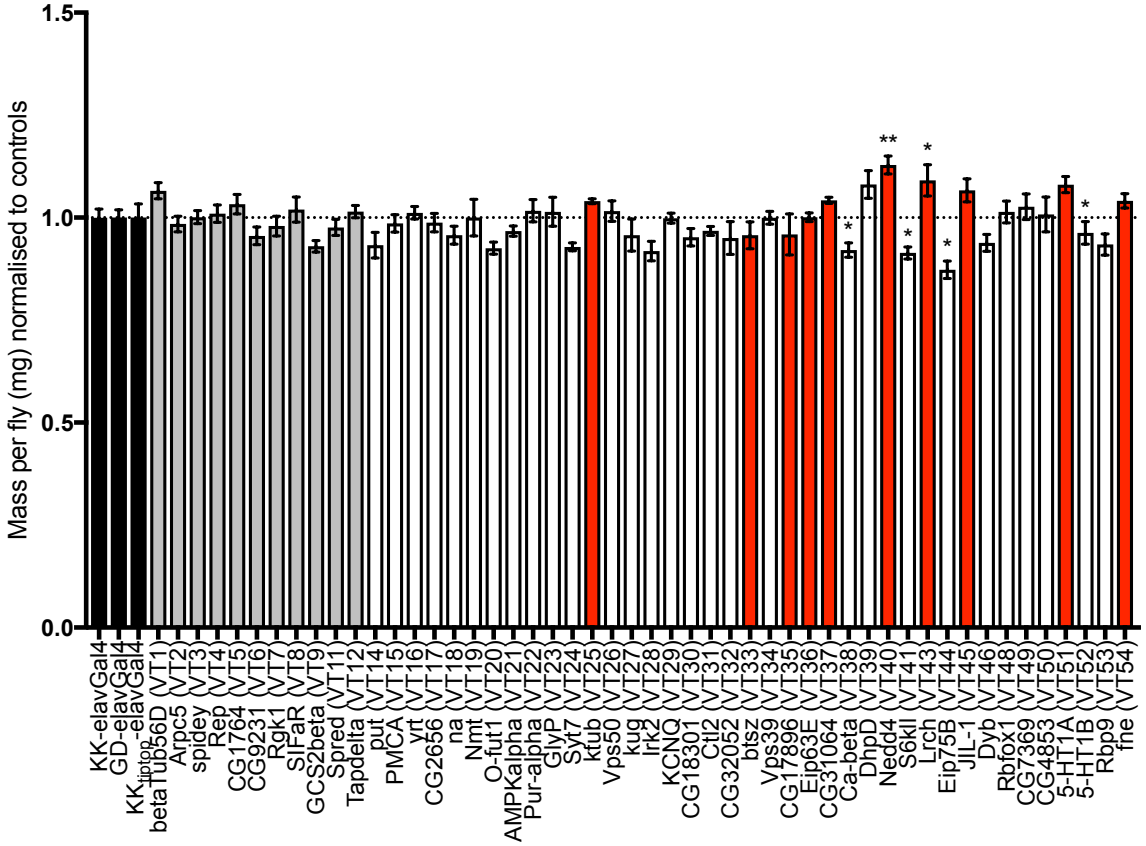


Fig. 6.8 Mass of flies with neuronal RNAi of transcriptomics-identified genes. Background control lines are black, orexigenic genes are grey, anorexigenic genes are white, and the 10 top hit genes are highlighted in red. Mean ± SEM is plotted of 5 repeats each containing 15 flies. Results were compared to appropriate background by Student's t-test.
* $p<0.05$, ** $p<0.01$

Glucose levels

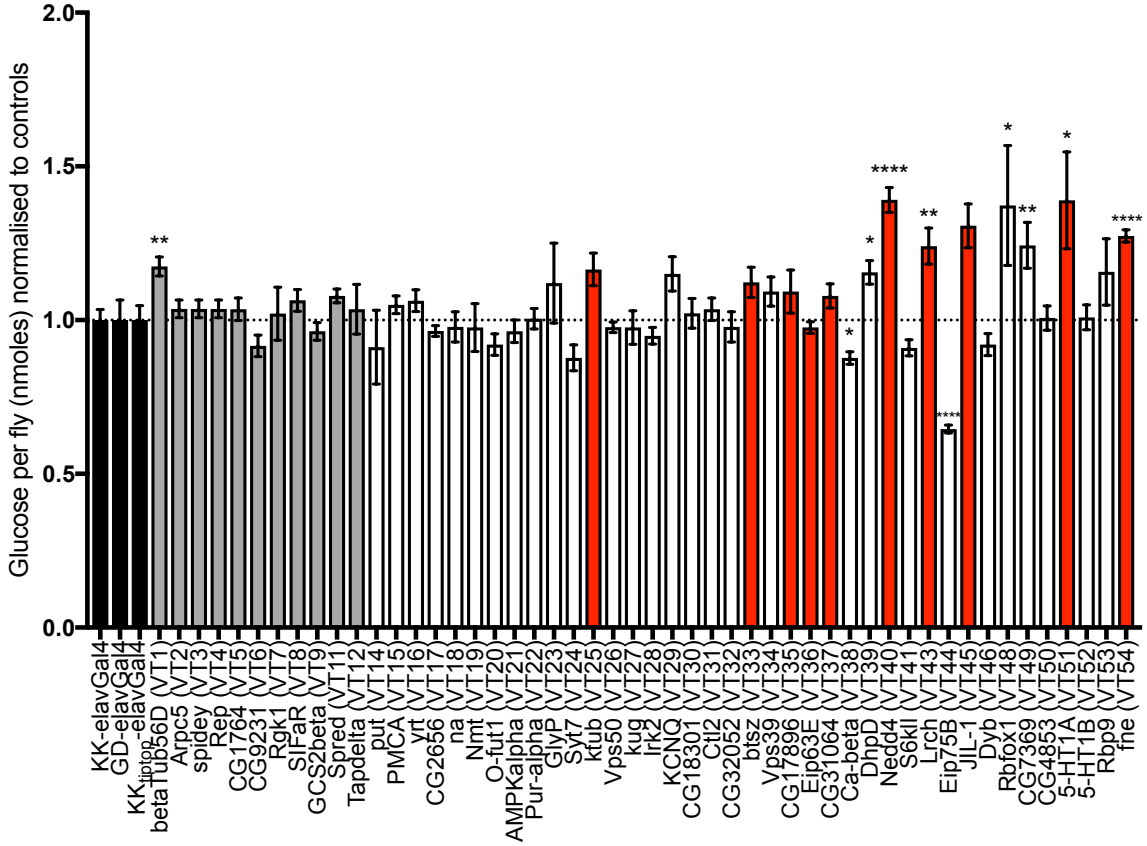


Fig. 6.9 Glucose levels in flies with neuronal RNAi of transcriptomics-identified genes. Background control lines are black, orexigenic genes are grey, anorexigenic genes are white, and the 10 top hit genes are highlighted in red. Mean ± SEM is plotted of 5 repeats each containing 15 flies. Results were compared to appropriate background by Student's t-test. * p<0.05, ** p<0.01, *** p<0.001, **** p<0.0001

Results summary

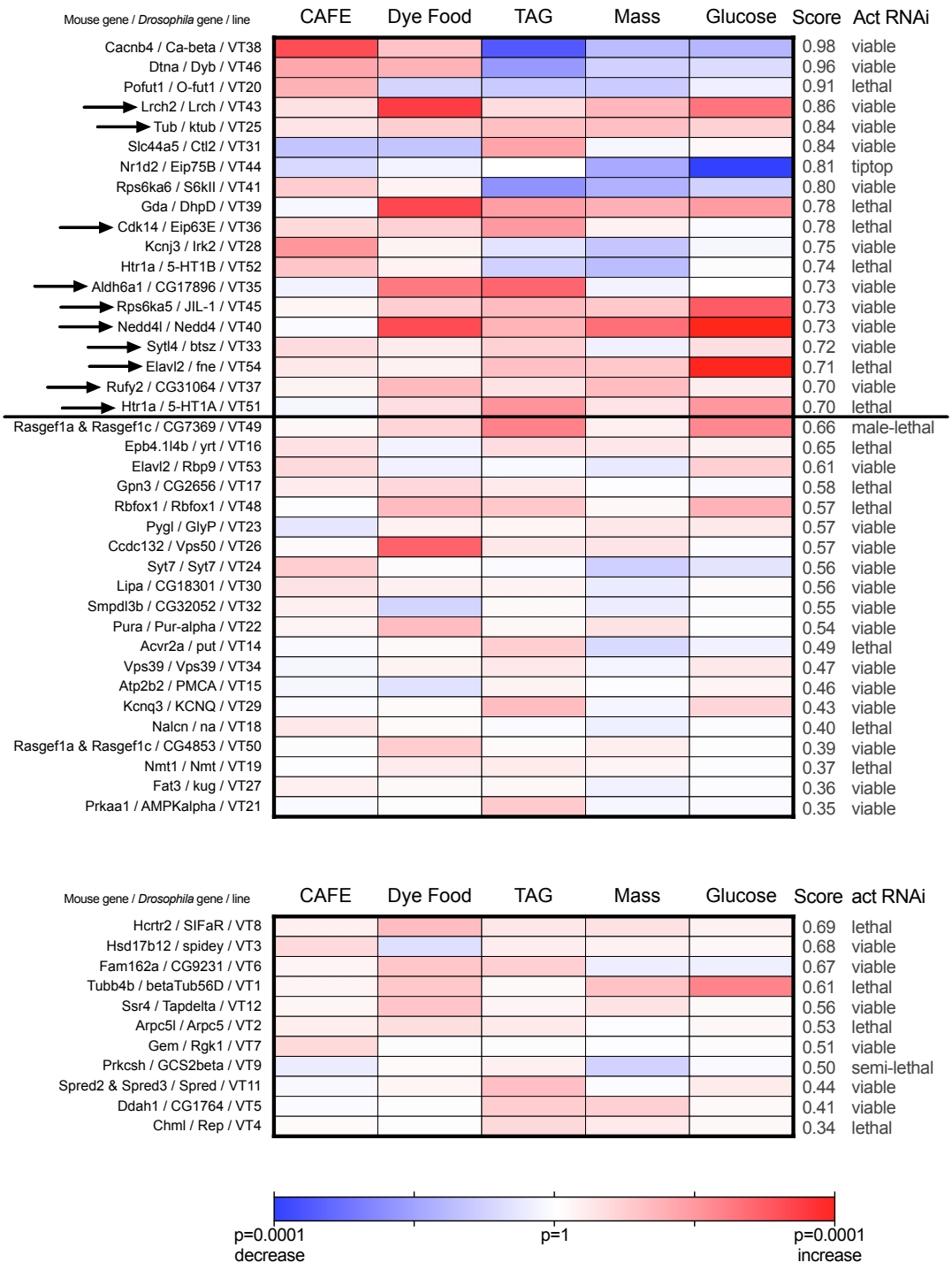


Fig. 6.10 RNAi screen of (top) anorexigenic genes and (bottom) orexigenic genes. Genes are ordered according to score. The top hit genes are highlighted with an arrow.

6.9 Published data regarding top hits

According to data available at Uniprot (Bateman et al., 2017), the top hit genes have a range of (predicted) biochemical functions, as summarised below:

Mouse Gene	Predicted Function (Uniprot)
Cdk14	cell cycle
Rufy2	endocytosis
Sytl4	protein transport
Nedd4l	protein regulation
Elavl2	translation regulation
Rps6ka5	transcription regulation
Htr1a	receptor
Lrch2	signal transduction
Tub	signal transduction
Aldh6a1	valine and pyrimidine metabolism

Of the 10 top hit genes, there are published studies for 6 of the genes (Tub, Htr1a, Rufy2, Sytl4, Cdk14, Nedd4l) in which an energy homeostasis phenotype was seen when the gene is mutated in mouse and/or humans. No such data was found for Aldh6a1, but the gene has been statistically associated with metabolic disease (Timmons et al., 2018). Finally, for Lrch2, Elavl2, and Rps6ka5 no relevant studies were found. This data is summarised in table 6.2.

Also, it is worth noting that there are published studies linking several of these genes to the same human diseases that were used in chapter 3 as negative genetic controls, for example ELAVL2 and schizophrenia (Yamada et al., 2011), and SYTL4 and autism (Butler et al., 2015).

Table 6.2 Published energy homeostasis phenotypes of the top hits from the *Drosophila* screen

Mouse gene	Mouse KO phenotype	Reference	Human data	Reference
Tub	Increased body weight Increased insulin Increased body fat Increased food intake	Noben-Trauth et al. (1996)	TUB expression negatively correlated with body weight and obesity	Nies et al. (2018)
Htr1a	Increased fat mass Hypoactivity Altered glucose response	Dickinson et al. (2016)		
Rufy2	Decreased insulin Altered glucose response	Dickinson et al. (2016)		
Sytl4	Decreased body weight Increased insulin	Jackson Laboratory (2018)		
Cdk14	Decreased fat	Dickinson et al. (2016)		
Nedd4l	Semi-lethal Many detrimental Haploinsufficiency protects from HFD-induced obesity	Dickinson et al. (2016) Jackson Laboratory (2018) Li et al. (2015)		
Aldh6a1	No relevant data found		Associated with metabolic disease	Timmons et al. (2018)
Lrch2	No relevant data found			
Elavl2	No relevant data found			
Rps6ka5	No relevant data found			

6.10 Discussion

6.10.1 Selection of mouse genes for study

The RNAseq data provided from Henry et al.'s paper reported the expression levels of 4889 genes whose expression was significantly different between the fed and fasted states in POMC and/or NPY neurons. Many different criteria could have been used to further filter the data but the aim here was to identify some genes which are promising candidates for further study, and not to provide a definitive answer as to the role of each gene. Therefore, there is no single "correct" set of filters. In this study, mouse genes were chosen if they showed (i) significant changes in expression between the fed and fasted states, (ii) reciprocal changes in expression between POMC and NPY neurons, and (iii) a strong homologue in humans, because genes which meet all of these three criteria are the most likely to have future therapeutic potential for obesity (section 6.4).

6.10.2 Identification of *Drosophila* gene orthologues

Drosophila orthologues of the selected hypothalamic fasting mouse genes were identified using the ENSEMBL ortholog tool. Using this method, 58% of the mouse genes were found to have an orthologue in flies (section 6.4), which is approximately the same as the genome-wide average of 53% (Wangler et al., 2017). This level of conservation was high enough for the data set to benefit from a *Drosophila*-based screen.

Given that energy homeostasis and feeding are important fundamental processes for all organisms, it is likely that any associated genetic data will show either similar or increased levels of conservation between species, and thus could be interrogated using the same *Drosophila*-based approach. The absence of homologues for 42% of the hypothalamic RNAseq genes is not a problem in this particular scenario, because the aim was to select promising candidates from a large list. However, if a data set requires direct comparison of the genes (for example, discriminating between genes in the same GWAS locus) but there is not increased levels of conservation of genes, then a *Drosophila*-based screen may not be appropriate.

The fact that the percentage of the hypothalamic fasting genes with an orthologue in *Drosophila* was close to the genome-wide average highlights the importance of the GWAS genes which were shown to have greatly increased levels of conservation — 79% of the human genes had a fly orthologue (see section 4.14.1).

The ENSEMBL orthologue tool used here searches for homologous genes based mostly on sequence similarity. The screening strategy used here relies on the assumption that, as well as conservation of sequence, genes also have conserved functions in *Drosophila* and cause similar phenotypes when perturbed. This is not always the case, and evolution has repurposed many conserved molecular systems such as signal transduction pathways (Wangler et al., 2017). This is a major limitation of working with *Drosophila* which is only distantly related in evolution to humans. However, as mentioned previously, the aim of this screen was to find top candidates for further study from the large list of hypothalamic transcriptomic genes, and not to give a definitive answer as to the role of each gene. Therefore, any genes which are detected by the *Drosophila* screen are much more likely to be important in humans because evolution has conserved their function in energy homeostasis.

The genes studied in this chapter were all chosen because they are expressed in the mouse hypothalamus. To increase the likelihood of the *Drosophila* genes being functional homologues, they too had to be expressed in the fly CNS. According to published microarray data (Graveley et al., 2011) only 67% of the genes were expressed in the *Drosophila* CNS. As discussed in chapter 4.14.1, this microarray data may under-estimate how many genes display conserved expression. By comparison, 92% of the GWAS BMI genes were expressed in the *Drosophila* CNS (chapter 4.6) which again highlights the importance of that set of genes.

As explored in chapters 3-5, RNAi seemed the most appropriate model for use in the *Drosophila*-based screen for a number of reasons. However, for the gene set in this chapter, some genes did not have any siRNA line available, and others only had siRNA lines with multiple predicted off-target effects. In total siRNA lines were only ordered for 82% of the *Drosophila* genes of interest. Therefore, although the number of mouse genes with fly homologues was calculated above, this is a slight over-estimate of the number of genes which can actually be studied. Again, this is not a problem in this particular scenario but is an important factor to consider for any future applications of this screen.

Therefore, in total, 52 mouse genes were studied in *Drosophila*.

6.10.3 A large-scale high-throughput *in vivo* screen of genes

The primary aim of this thesis was to develop a high-throughput *in vivo* screen suitable for use on large data sets for feeding behaviour and energy homeostasis phenotypes. The speed for which the data in this chapter was collected (6 weeks, 52 genes of interest, 5 phenotypes each, section 6.5) shows that the developed screen undeniably fulfils these criteria.

In the same time period it would not be possible to breed a mouse to adulthood. Further, mice with perturbation (especially tissue-specific perturbation) do not yet exist for most genes and so would have to be created *de novo*. Therefore, the overall aim of this thesis was achieved. The only caveat is that not all of the genes could be studied in *Drosophila* due to lack of a homologue, but all of the genes are present in mice.

6.10.4 Phenotype of genes

Each gene was phenotyped using neuronal RNAi and the following five assays: CAFE assay food intake, ingestion of dye food following fasting, TAG levels, wet mass, and glucose levels (section 6.8). Further, whether the gene was lethal or viable with whole body RNAi was assessed.

Lethality

The proportion of genes which were lethal with universal RNAi (31%) is close to the genome-wide average (section 6.6). This tentatively suggests that it would be possible to create viable mouse KO models. Further, this result highlights the importance of the GWAS genes for which the lethality was found to be increased (section 4.7).

Over-expression of *tiptop*

Whenever possible, siRNA lines with a GD background were ordered in preference to KK. However, there were still 17 genes for which KK lines had to be used. Of these, 7 were lethal when crossed to act-GAL4 so were considered as possibly being affected by *tiptop*, thus hampering the interpretation of the data and reducing confidence in the results for these 7 genes (section 6.7). Therefore, quite a significant proportion of genes were difficult to study due to *tiptop* ambiguity. Although it is better to have a false negative result than a false positive for the purpose of candidate gene selection this factor does decrease the proportion of genes which can be studied.

However, just 10% of genes are only available as KK lines (Dietzl et al., 2007), and only 25% of KK lines are affected by *tiptop* (Vissers et al., 2016). Therefore, it is only (approximately) 2.5% of all *Drosophila* genes for which the only option is a *tiptop*-affected siRNA line. This is much lower than the 25% of the genome which has a lethal phenotype when mutant (Wangler et al., 2017) and so perhaps the impact of this problem will be negligible in most screens. Indeed, other published papers (for example Pendse et al. (2013) and Baranski et al. (2018)) did not state any control for *tiptop* over-expression. Therefore, this limitation of the RNAi model may not be hugely detrimental, but should always be taken into consideration.

Anorexigenic vs orexigenic genes

There were multiple noteworthy results of interest in the anorexigenic genes (section 6.8). It is not surprising that significant increases in feeding behaviour are seen after RNAi of these genes given that:

1. activation of NPY neurons triggers feeding behaviours (Aponte et al., 2011)
2. fasting activates NPY neurons (Hahn et al., 1998)
3. fasting decreases expression of the anorexigenic genes in NPY neurons (Henry et al., 2015)
4. and RNAi causes a decrease in expression of these genes in fly neurons.

In contrast, there are no results of note in the orexigenic genes (which decrease in POMC cells during fasting), based on either the score or the individual assay results. Perhaps increasing expression of these genes would produce a phenotype.

However it should be noted that more anorexigenic genes were studied than orexigenic genes, and so the fact that more significant phenotypes were observed in the former may not be significant.

Genes which were not hits

No hits were found in the *Drosophila*-based screen for 31 of the hypothalamic mouse genes. Possible reasons that these genes did not test positive in the *Drosophila* screen are the same as those discussed in chapter 4 in relation to the GWAS genes. Briefly: the siRNA was not effective; the *Drosophila* gene is not a functional orthologue of the mouse gene; the gene only causes a phenotype when up-regulated; the screen was too stringent; and/or it is difficult to ascertain the true phenotype of potential *tiptop*-affected lines. Alternatively, it is possible that these mouse genes do not actually play a role in energy homeostasis, and the changes in expression observed upon fasting are instead part of a different pathway.

Genes with discordant phenotypes

The homologues of Cacnb4, Dtna, Pofut1 and Rps6ka6 showed an increase in food intake but a decrease in measures of physiology. The homologues of Slc44a5 and Kcnj3 showed the opposite phenotype. It is possible that these genes also had an energy expenditure phenotype which could explain the discordant results, but which was not measured as part of the screen. Indeed, CACNB4 mutations have been shown to cause ataxia and epileptic seizures in humans (Escayg et al., 2000) and deletion of Pofut1 in mice affects skeletal muscle and motor neurons (Zygmunt et al., 2017). These genes are perhaps worthy of further investigation in mammalian models, but should not be prioritised.

Gene with decrease phenotypes

The homologue of Nr1d2 showed significant decreases in both energy intake and physiology. As mentioned previously, it is possible that this genetic perturbation made the flies unhealthy, rather than caused an energy homeostasis phenotype. In agreement, mice with KO of Nr1d2 show many pathological phenotypes across many different organs and systems (Dickinson et al., 2016).

Genes which were hits

Finally, there are 10 genes for which neuronal RNAi in *Drosophila* causes an increase in both food intake and measures of physiology. It is these genes which should be considered for further study in complex mammalian models such as mice. This high hit rate (21% of the

genes studied) is unsurprising, given that the starting list of genes was not random but was instead selected based on a suspected role in energy homeostasis.

For 7 of these genes, there is published data from mouse and/or human studies. For Tub, Htr1a, and Nedd4l there is strong evidence for a role of the gene in energy homeostasis, and for Rufy2, Sytl4, Cdk14 and Aldh6a1 there is suggestive evidence. The supportive evidence demonstrates the utility of using *Drosophila* as a high-throughput screen of genes for energy homeostasis phenotypes.

For Lrch2, Elavl2, Rps6ka5, no published data was found regarding energy homeostasis phenotypes. Although this data will, one day, be available, the speed of results generation in mice is exponentially slower than that for *Drosophila*. Therefore, at least for the time being, screens such as the one used here in *Drosophila* are the best way to obtain energy homeostasis information for a large number of genes, so allowing prioritisation of which mouse models should be made.

The top hit genes are not the most highly expressed genes in the mouse neurons, nor the ones which show the largest changes in expression, and they have a variety of different functions. Therefore, it is unlikely that these genes would have been selected for further study based solely on the transcriptomic data. This highlights an important advantage of the *Drosophila*-based screen: its ability to assess multiple candidates in an unbiased manner (the screen was done blinded to genotype) allowing discoveries that may otherwise go un-investigated.

6.10.5 Further work

The *Drosophila*-based screen in this chapter identified 10 genes which are worthy of further investigation. These genes have a variety of predicted functions and thus so require a multi-disciplinary approach to study their underlying biochemistry and how this impacts energy homeostasis.

On a cellular level, many experiments could be envisioned to answer questions such as what other genes/gene products/substrates do they act on? By which genes/proteins are they themselves regulated? *Drosophila* may prove to be a good model here as individual pathways are often conserved even in cases where the overall function is not (Wangler et al., 2017). Further, where in the cell do the gene products act? And what is their molecular function? These questions are perhaps easiest to answer using cell culture techniques.

On a physiological scale, it is already known that these genes respond to starvation, but are they affected by other nutritional perturbations such as HFD? Or by metabolic hormones such as leptin, insulin or ghrelin? Do mice with KO of these genes eat more and/or have increased stores of energy? Importantly for drug development, are mice with whole-body perturbations of the gene viable and healthy? Do humans with KO of these genes exist, and if so, do they have any measurable phenotype? Perhaps the most informative experiments would to look at mice with hypothalamic-specific perturbation of the genes. Unlike in *Drosophila*, there are no readily available lines for doing this in mice and the experiment would require injection of the mutating reagents (for example Cas9 plus gRNAs) into the developing mouse hypothalamus which is a difficult, expensive and time-consuming experiment, so would not be wise to carry out on a gene for which there is no other data suggesting that the results would be profitable.

6.11 Summary

In summary, in this chapter, a *Drosophila*-based screen has proved itself to be a useful tool for selecting top candidate genes for a role in energy homeostasis from a large genetic data set. A set of 52 genes were identified which, based on hypothalamic transcriptomic data from Henry et al. (2015), respond to fasting and thus may play a role in energy homeostasis. *Drosophila* homologues of these genes were identified and then phenotyped using the screen developed in chapters 3 and 4. The results were used to identify top candidates for further investigation in mice to advance our understanding of how the CNS controls energy homeostasis, and thus aid development of therapeutics for obesity.

Chapter 7

Discussion and Conclusions

The causes of obesity are complex with biological, social, and psychological factors all playing a role. It has been shown that genetics has a strong influence on obesity risk, and understanding the underlying biology may provide a new avenue into developing novel pharmaceutical reagents which are both specific and effective for the treatment of obesity. The advent of modern genomics techniques is massively expanding the number of genes which are potentially involved at a rate with which functional validation studies in mice — the traditional model organism of choice for the study of energy homeostasis — cannot keep pace.

Development and use of a high-throughput screen is important if we are to be able to fully understand the genetic contribution to obesity. In this light, the aim of this thesis was to develop a high-throughput screen for genes that are involved in the neuronal control of energy homeostasis and feeding behaviour using *Drosophila melanogaster* as a model organism, and to use this screen to explore data sets from GWAS and transcriptomics.

It is important to note that the primary goal here is not to add to knowledge about *Drosophila* biology, as a plethora of such data already exists. Instead, the goal is to use the fly model as an efficient screen to help move mammalian statistical data towards a functional understanding. As described in chapter 1, studies with similar objectives have been done before, but this is a novel approach due to the combination of: use of multiple assays to look at several related characteristics; a focus on feeding behaviour; and starting from a list of mammalian genes.

Development of a *Drosophila*-based screen

A *Drosophila*-based screen for energy homeostasis phenotypes was developed and refined through chapters 3, 4, 5 and 6. To begin, this involved finding and adapting assays from published literature. These assays were tested for their ability to detect metabolic perturbation, their ability to distinguish negative control genes from positive control genes, and whether they were high-throughput. Based on these criteria, not all assays were used in the final screen: the dry mass was found to be too insensitive, and the dye food absorption assay was shown to produce inaccurate results. However, the CAFE, dye food over-feeding, wet mass, TAG, glucose, starvation, climbing and larval feeding assays were all found to be suitable to use together to form the screen. The rate-limiting step was the CAFE assay due to possession of only a limited amount of equipment (the lids needed to hold the feeding capillary tubes) but this could easily be remedied by purchase of more. In addition, an assay to measure feeding motivation in adults was developed, although this was low-throughput and so only suitable for use on the top hit genes.

As stated above, one way in which the screen was validated involved control genes. However, as was discussed in chapters 3, 4 and 6, it is possible that the genes chosen to use as both positive and negative controls were not ideally suited for this purpose. In particular, the negative controls were genes which have been associated with other diseases, but these diseases are not independent of weight, and it was later found that many of the BMI GWAS genes and hypothalamic transcriptomics genes that were studied have been related to these same diseases. Therefore, it is possible that the screen is overly stringent and will miss significant phenotypes. However, given that the purpose of the screen was to select top candidate genes, it is highly preferable for the screen to have false negative results than for it to to be too lenient and detect false positives.

Development of the screen also involved comparing and contrasting three different *Drosophila* genetic models: LoF (mostly P-element mediated), CRISPR, and UAS_G-GAL4 RNAi. The various pros and cons of each method were discussed in chapters 3 and 5, and ultimately RNAi was found to be the most appropriate for use in this screen.

As was seen in chapter 6, the final screen is high-throughput, especially when compared to mice, and thus the primary aim of this thesis was met.

Limitations of a *Drosophila*-based screen

The *Drosophila*-based screen was used to generate novel results in chapters 4 and 6, thus proving its utility. However, there are a few limitations of both the screen itself and the reliability of the data obtained.

It is possible that the siRNAs in some of the lines had off-target effects which were responsible for the phenotype observed, rather than the gene of interest. Thus, results should be viewed with some degree of caution. However, since the studied genes were selected because they are already suspected to be involved in energy homeostasis, the chance of an off-target effect causing the same phenotype is less likely.

Some might argue that the small size of *Drosophila* is a disadvantage, since it makes phenotyping more difficult due to the requirement for more sensitive equipment. However, this can be easily rectified by measurements of groups of flies, and is compensated for by the fact that their small size is one of the reasons that working with *Drosophila* is so economical and high-throughput.

Instead, the major limitation of using *Drosophila* to screen the human/mouse genes is that not every gene can be studied. This may explain why some of the genes which are already known to have a role in energy homeostasis (for example TMEM18) were not detected in the screen. Genes are lost at five stages:

1. A gene of interest may not have an identifiable sequence homologue in *Drosophila*. The number of genes for which an orthologue was found was increased compared to the genome average for the GWAS BMI genes (chapter 4), but not for the hypothalamic transcriptomic genes (chapter 6). Since human genes linked to monogenic disorders also show increased levels of conservation in *Drosophila* (Fortini et al., 2000), this suggests that genetic data linked directly to diseases (like GWAS data) could gain the most benefit from the *Drosophila*-based screen in the future.

2. Any sequence homologues which are identified may not also be a functional homologue of the human gene, potentially leading to false negative results. Whether this is a problem depends on the exact purpose of the specific screen. For example, with the hypothalamic transcriptomics data in chapter 6, the purpose was to choose genes for further study, and there is no single "correct" answer and so a false negative result is allowable. However, when used for discriminating between genes in the same GWAS locus as in chapter 4, it is only a fair test if all of the genes can be studied.

3. 9% of *Drosophila* genes do not have any UAS_G-siRNA lines available to order, and for some genes the only siRNA lines available have multiple known off-target effects. Again this may not be a problem depending on the purpose of the study. If it is important to study all of the genes, then another model could be used, for example LoF flies, although there is no guarantee that all of the genes of interest will be available with any model.

4. Another limiting factor is lethality. No genes in this thesis were found to be lethal with neuronal RNAi, but this is unlikely to be universally true. Further, a different GAL4 driver may increase the rates of lethality, as was found in the paper by Baranski et al. (2018).

5. Finally, for 2.5% of genes, the only siRNA line available is affected by over-expression of *tiptop* which causes strong energy homeostasis phenotypes. However, there is currently no definitive list of which lines are affected, meaning that this genotype had to be ascertained instead by looking at the phenotype of the flies. Therefore, there is a large degree of uncertainty regarding the "true" phenotype of these lines, meaning that false negative results may be introduced. Further, it meant that GD lines were always chosen in preference, which the VDRC stock centre states achieve less efficient RNAi compared to the KK lines (VDRC, 2016), potentially introducing more false negative results.

Advantages of a *Drosophila*-based screen

The major advantage of the *Drosophila*-based screen developed and used in this thesis is the speed with which results are generated. In chapter 6, 50 genes were screened in 6 weeks. The speed could be increased further by purchase of more CAFE assay lids. By comparison, this figure is completely unrivalled by mice: the IMPC, whose stated aim is to screen mice in a high-throughput manner, have a large number of people working on an optimised production pipeline, but only hope to generate 100 KO mouse lines per year at a cost of £28,000 per gene (Moore, 2010). Even if a mutant mouse line already exists for a gene of interest, it takes more than six weeks just to breed new mice for experiments. Further, the IMPC mice have whole-body KO mutations of the genes, which as was seen repeatedly through this thesis, is often detrimental and the information provided is not as useful as neuron-specific data. Although neuron-specific experiments are possible in mice, there are no readily available lines for doing so and so the mouse models must be produced *de novo*.

In both of the data sets screened in this thesis (BMI GWAS in chapter 4 and hypothalamic transcriptomics in chapter 6), there were a large number of positive findings. This suggests that much of the biological machinery regulating energy homeostasis is conserved in evolution, which is perhaps unsurprising given its fundamental importance for life. Therefore there is a strong argument for the utility of high-throughput functional genetic screens in simple model organisms such as *Drosophila*.

In its current set-up, the screen can be easily adapted to suit different data sets. Our lab group is interested in the neuronal control of feeding behaviour, but the screen could be applied to other tissues. For example, GWAS has identified loci which affect waist-to-hip ratio, and it has been shown that genes in these loci are preferentially expressed in adipose tissue. Thus it may be pertinent to explore this data set using RNAi in the *Drosophila* fat body driven by Lsp2-GAL4 (Lazareva et al., 2007).

Findings from two *Drosophila*-based screens of mammalian genetic data

The developed and validated screen was used to investigate two mammalian genetic data sets related to energy homeostasis. 53 human genes in loci that GWAS studies have associated with BMI were explored in chapter 4. Then, in chapter 6, 52 mouse genes which have been shown to be reciprocally regulated in POMC and NPY hypothalamic neurons by fasting were studied.

For both data sets, the screen found several genes with a positive energy homeostasis phenotype in *Drosophila*. For approximately half of these genes, there are published studies in mouse and/or humans which show that they do indeed play a role in energy homeostasis: PCSK1, NEGR1, MTIF3, CDKAL1, and BDNF from the human GWAS BMI genes, and Tub, Htr1a, Rufy2, Sytl4, Cdk14, Nedd4l, and Aldh6a1 from the mouse hypothalamic transcriptomics genes. This provides validation that the screen is capable of detecting genes with a relevant phenotype.

The most interesting genes highlighted by the screens were those for which a positive energy homeostasis phenotype was observed but for which no mouse/human data has yet been published. It is these genes which have the most potential to be rewarding for a lab to study further. From the screen of human GWAS data these genes were: LRRN6C, KLF9, NRXN3, GABRG1, SEC16B, TUFM, NUDT3, and COL4A3BP. From the screen of mouse hypothalamic transcriptomics data these genes were: Lrch2, Elavl2, and Rps6ka5. Thus, the

second aim of this thesis was met: to find novel evidence for a role of genes-of-interest in energy homeostasis.

Further, RNAi of some of the genes in the screen caused the flies to have a negative energy homeostasis phenotype. It is not possible to distinguish the cause of such results between (i) a true lean phenotype caused directly by the gene of interest acting on feeding pathways, or (ii) a side-effect of the gene perturbation which makes the flies unhealthy and therefore less developed and/or less motivated to eat. However, a very unhealthy phenotype has also been reported in mouse models and human patients for some of these genes, suggesting that the latter hypothesis is true: SLC39A8, MTCH2, RPL27A, Nr1d2. Therefore an important use of a *Drosophila*-based screen could be to filter out any genes which are detrimental (or lethal) before resources are used to develop a mouse model which turns out ultimately to just be unhealthy or non-viable.

There is currently debate as to whether GWAS studies of BMI are useful in relation to the study of obesity (see Müller et al. (2018) and Speakman et al. (2018)). Two results from this thesis highlighted how significant the genes which have been identified by GWAS to date are: a higher proportion of the genes are conserved in *Drosophila* compared to the genome average, and a higher proportion are lethal with whole body perturbation, both of which suggest that the genes play an important role in physiology. In comparison, both of these figures are close to the genome average for the hypothalamic transcriptomic genes.

Future work

Drosophila studies of top hits

Through this thesis neuron-specific RNAi was found to be the most suitable model for the purpose of a screen for genes involved in the neuronal control of energy homeostasis. However, RNAi does have the potential for off-target effects. Given the investment required to follow up a gene in a mouse model, it is perhaps wise to first confirm the results seen in another *Drosophila* line. The easiest method of doing so would be to use a different siRNA line. However, for many genes this does not exist, and for those which it does this second line is equally likely to have off-target effects, or may contain a non-functional siRNA. A more accurate approach may be to develop a functional method of performing tissue-specific CRISPR, and various approaches for trying to do so were discussed in chapter 5. However, the genes studied all originally came from lists of human/mouse data for which

there is already statistical data suggesting a role in energy homeostasis, and so perhaps further validation in *Drosophila* is unnecessary.

Mouse studies of top hits

The next step would be to study in mice the genes which were top hits in the *Drosophila*-based screen. Based on the *Drosophila* work throughout this thesis, it is perhaps worth investing in the creation of brain-specific perturbations of the genes in mice. Examples of potential experiments were put forward in chapter 6 to investigate both the biochemical function of these genes as well as phenotypes caused by their perturbation. If mouse studies prove successful, the final goal would be studies in human, ultimately leading to the development of therapeutics for obesity.

Future screens in *Drosophila*

The same *Drosophila* screen developed and used in this thesis could also be used to explore other large mammalian genetic data sets. As just one example, the lab group has looked at transcriptional changes in different parts of the mouse hypothalamus (in contrast to Henry *et al.* who looked at different cells but spread across the whole hypothalamus) in all of the fed, fasted and HFD-fed states. There are also many published papers which present transcriptomic data in various different tissues and/or states, and multiple GWASs for phenotypes related to energy homeostasis including more recent studies of BMI (Locke et al., 2015; Akiyama et al., 2017).

As discussed above, the screen can be adapted to look at other tissues. It can also be adapted to look at different/more phenotypes. In particular, a more sensitive way of measuring movement and/or energy expenditure would be a useful addition. One solution would be to use the same climbing assay used in chapters 3 and 4, but to increase the climbing distance, for example some labs put single flies into a 50cm long tubes. Alternatively, a method which simultaneously and automatically measures both food intake and movement of a fly is currently in development in other labs (Ja, 2018). Other work in the lab group is currently developing a respirometer to measure production of carbon dioxide by flies as a measure of metabolic rate. However, all three of these methods are low-throughput, and thus not suitable for studying large data sets. There are also more assays of feeding behaviour which could be added to the screen, such as larval mouth hook contraction (Bhatt and Neckameyer, 2013), adult proboscis extension (Shiraiwa and Carlson, 2007) and food preference (Bantel and Tessier, 2016), but again these are all low-throughput. Although not suitable for large scale

screens, these slower assays could be used as a secondary screen of just the top hit genes as further verification of their phenotype.

Conclusions

In conclusion, in this thesis a high-throughput screen was developed in *Drosophila* for use in functional validation of statistical genomic data from humans/mice. This screen was used to investigate two data sets (GWAS BMI genes and fasting hypothalamic transcriptomics genes) and identified many candidates worthy of further investigation. Some of these were already known which provides important validation of the screen, but many are uninvestigated and thus provide novel results. Follow-up studies of these genes in mammalian models may, one day, progress our understanding of, and thus treatment of, obesity.

References

Abel, E. D., Peroni, O., Kim, J. K., Kim, Y. B., Boss, O., Hadro, E., Minnemann, T., Shulman, G. I., and Kahn, B. B. (2001). Adipose-selective targeting of the GLUT4 gene impairs insulin action in muscle and liver. *Nature*, 409(6821):729–33.

Achapu, M., Pompei, P., Polidori, C., de Caro, G., and Massi, M. (1992). Central effects of neuropeptide K on water and food intake in the rat. *Brain research bulletin*, 28(2):299–303.

Akiyama, M., Okada, Y., Kanai, M., Takahashi, A., Momozawa, Y., Ikeda, M., Iwata, N., Ikegawa, S., Hirata, M., Matsuda, K., Iwasaki, M., Yamaji, T., Sawada, N., Hachiya, T., Tanno, K., Shimizu, A., Hozawa, A., Minegishi, N., Tsugane, S., Yamamoto, M., Kubo, M., and Kamatani, Y. (2017). Genome-wide association study identifies 112 new loci for body mass index in the Japanese population. *Nature genetics*, 49(10):1458–1467.

Al-Anzi, B. (2010). The leucokinin pathway and its neurons regulate meal size in Drosophila. *Curr Biol*, 20(11):969–978.

Al-Saffar, Z., Grainger, J., and Aldrich, J. (1996). Temperature and humidity affecting development, survival and weight loss of the pupal stage of Drosophila melanogaster, and the influence of alternating temperature on the larvae. *Journal of Thermal Biology*, 21(5-6):389–396.

Albin, S., Kaun, K., Knapp, J.-M., Chung, P., Heberlein, U., and Simpson, J. (2015). A Subset of Serotonergic Neurons Evokes Hunger in Adult Drosophila. *Current Biology*, 25(18):2435–2440.

Alsters, S. I. M., Goldstone, A. P., Buxton, J. L., Zekavati, A., Sosinsky, A., Yiorkas, A. M., Holder, S., Klaber, R. E., Bridges, N., van Haelst, M. M., le Roux, C. W., Walley, A. J., Walters, R. G., Mueller, M., and Blakemore, A. I. F. (2015). Truncating Homozygous Mutation of Carboxypeptidase E (CPE) in a Morbidly Obese Female with Type 2 Diabetes Mellitus, Intellectual Disability and Hypogonadotrophic Hypogonadism. *PloS one*, 10(6):e0131417.

Altschul, S., Gish, W., Miller, W., Myers, E., and Lipman, D. (1990). Basic local alignment search tool. *J. Mol. Biol.*, 215:403–410.

Aponte, Y., Atasoy, D., and Sternson, S. M. (2011). AGRP neurons are sufficient to orchestrate feeding behavior rapidly and without training. *Nature Neuroscience*, 14(3):351–355.

Arrese, E. L., Flowers, M. T., Gazard, J. L., and Wells, M. A. (1999). Calcium and cAMP are second messengers in the adipokinetic hormone-induced lipolysis of triacylglycerols in Manduca sexta fat body. *Journal of lipid research*, 40(3):556–564.

Arrese, E. L., Rivera, L., Hamada, M., Mirza, S., Hartson, S. D., Weintraub, S., and Soulages, J. L. (2008). Function and structure of lipid storage droplet protein 1 studied in lipoprotein complexes. *Archives of biochemistry and biophysics*, 473(1):42–47.

Arrese, E. L. and Soulages, J. L. (2010). INSECT FAT BODY: ENERGY, METABOLISM, AND REGULATION. *Annual review of entomology*, 55:207–225.

Atkinson, M. A. and Maclaren, N. K. (1994). The pathogenesis of insulin-dependent diabetes mellitus. *The New England journal of medicine*, 331(21):1428–1436.

Attrill, H., Falls, K., Goodman, J., Millburn, G., Antonazzo, G., Rey, A., Marygold, S., and Flybase Consortium (2016). FlyBase: establishing a Gene Group resource for Drosophila melanogaster. *Nucleic Acids Res.*, pages 44(D1):D786–D792.

Austin, C., Lo, G., Nandha, K. A., Meleagros, L., and Bloom, S. R. (1995). Cloning and characterization of the cDNA encoding the human neuromedin U (NmU) precursor: NmU expression in the human gastrointestinal tract. *Journal of molecular endocrinology*, 14(2):157–169.

Ayers, K. L., Glicksberg, B. S., Garfield, A. S., Longerich, S., White, J. A., Yang, P., Du, L., Chittenden, T. W., Gulcher, J. R., Roy, S., Fiedorek, F., Gottesdiener, K., Cohen, S., North, K. E., Schadt, E. E., Li, S. D., Chen, R., and Van der Ploeg, L. H. T. (2018). Melanocortin 4 Receptor Pathway Dysfunction in Obesity: Patient Stratification Aimed at MC4R Agonist Treatment. *The Journal of Clinical Endocrinology & Metabolism*, 103(7):2601–2612.

Azevedo, F. A. C., Carvalho, L. R. B., Grinberg, L. T., Farfel, J. M., Ferretti, R. E. L., Leite, R. E. P., Jacob Filho, W., Lent, R., and Herculano-Houzel, S. (2009). Equal numbers of neuronal and nonneuronal cells make the human brain an isometrically scaled-up primate brain. *The Journal of comparative neurology*, 513(5):532–541.

Bantel, A. P. and Tessier, C. R. (2016). Taste Preference Assay for Adult Drosophila. *Journal of Visualized Experiments*, (115).

Baranski, T. J., Kraja, A. T., Fink, J. L., Feitosa, M., Lenzini, P. A., Borecki, I. B., Liu, C.-T., Cupples, L. A., North, K. E., and Province, M. A. (2018). A high throughput, functional screen of human Body Mass Index GWAS loci using tissue-specific RNAi Drosophila Melanogaster crosses. *PLOS Genetics*, 14(4):e1007222.

Bartholomew, N. R., Burdett, J. M., VandenBrooks, J. M., Quinlan, M. C., and Call, G. B. (2015). Impaired climbing and flight behaviour in Drosophila melanogaster following carbon dioxide anaesthesia. *Scientific Reports*, 5(1):15298.

Bateman, A., Martin, M. J., O'Donovan, C., Magrane, M., Alpi, E., Antunes, R., Bely, B., Bingley, M., Bonilla, C., Britto, R., Bursteinas, B., Bye-A-Jee, H., Cowley, A., Silva, A. D., Giorgi, M. D., Dogan, T., Fazzini, F., Castro, L. G., Figueira, L., Garmiri, P., Georghiou, G., Gonzalez, D., Hatton-Ellis, E., Li, W., Liu, W., Lopez, R., Luo, J., Lussi, Y., MacDougall, A., Nightingale, A., Palka, B., Pichler, K., Poggioli, D., Pundir, S., Pureza, L., Qi, G.,

Renaux, A., Rosanoff, S., Saidi, R., Sawford, T., Shypitsyna, A., Speretta, E., Turner, E., Tyagi, N., Volynkin, V., Wardell, T., Warner, K., Watkins, X., Zaru, R., Zellner, H., Xenarios, I., Bougueleret, L., Bridge, A., Poux, S., Redaschi, N., Aimo, L., Argoud-Puy, G., Auchincloss, A., Axelsen, K., Bansal, P., Baratin, D., Blatter, M.-C., Boeckmann, B., Bolleman, J., Boutet, E., Breuza, L., Casal-Casas, C., de Castro, E., Coudert, E., Cuche, B., Doche, M., Dornevil, D., Duvaud, S., Estreicher, A., Famiglietti, L., Feuermann, M., Gasteiger, E., Gehant, S., Gerritsen, V., Gos, A., Gruaz-Gumowski, N., Hinz, U., Hulo, C., Jungo, F., Keller, G., Lara, V., Lemercier, P., Lieberherr, D., Lombardot, T., Martin, X., Masson, P., Morgat, A., Neto, T., Nouspikel, N., Paesano, S., Pedruzzi, I., Pilbout, S., Pozzato, M., Pruess, M., Rivoire, C., Roechert, B., Schneider, M., Sigrist, C., Sonesson, K., Staehli, S., Stutz, A., Sundaram, S., Tognolli, M., Verbregue, L., Veuthey, A.-L., Wu, C. H., Arighi, C. N., Arminski, L., Chen, C., Chen, Y., Garavelli, J. S., Huang, H., Laiho, K., McGarvey, P., Natale, D. A., Ross, K., Vinayaka, C. R., Wang, Q., Wang, Y., Yeh, L.-S., and Zhang, J. (2017). UniProt: the universal protein knowledgebase. *Nucleic Acids Research*, 45(D1):D158–D169.

Baumbach, J., Hummel, P., Bickmeyer, I., Kowalczyk, K., Frank, M., Knorr, K., Hildebrandt, A., Riedel, D., Jäckle, H., and Kühnlein, R. (2014). A Drosophila In Vivo Screen Identifies Store-Operated Calcium Entry as a Key Regulator of Adiposity. *Cell Metabolism*, 19(2):331–343.

Bechtold, D. A., Ivanov, T. R., and Luckman, S. M. (2009). Appetite-modifying actions of pro-neuromedin U-derived peptides. *American journal of physiology. Endocrinology and metabolism*, 297(2):E545–51.

Behrouz, B., Vilariño-Güell, C., Heckman, M. G., Soto-Ortolaza, A. I., Aasly, J. O., Sando, S., Lynch, T., Craig, D., Uitti, R. J., Wszolek, Z. K., Ross, O. A., and Farrer, M. J. (2010). Mitochondrial translation initiation factor 3 polymorphism and Parkinson's disease. *Neuroscience Letters*, 486(3):228–230.

Beller, M., Bulankina, A. V., Hsiao, H.-H., Urlaub, H., Jäckle, H., and Kühnlein, R. P. (2010). PERILIPIN-Dependent Control of Lipid Droplet Structure and Fat Storage in Drosophila. *Cell Metabolism*, 12(5):521–532.

Bertagna, X. (1994). Proopiomelanocortin-derived peptides. *Endocrinology and metabolism clinics of North America*, 23(3):467–485.

Beshel, J. (2016). Plus-size model: A neural circuit governing food cue perception and the making of an obese fly. In *EMBO Workshop on Neural Control of Metabolism and Eating Behavior*.

Beshel, J., Dubnau, J., and Zhong, Y. (2017). A Leptin Analog Locally Produced in the Brain Acts via a Conserved Neural Circuit to Modulate Obesity-Linked Behaviors in Drosophila. *Cell metabolism*, 25(1):208–217.

Beshel, J. and Zhong, Y. (2013). Graded Encoding of Food Odor Value in the Drosophila Brain. *The Journal of Neuroscience*, 33(40):15693–15704.

Bhatt, P. K. and Neckameyer, W. S. (2013). Functional Analysis of the Larval Feeding Circuit in Drosophila. *J Vis Exp*, (81):e51062.

Bi, J., Xiang, Y., Chen, H., Liu, Z., Gronke, S., Kuhnlein, R. P., and Huang, X. (2012). Opposite and redundant roles of the two Drosophila perilipins in lipid mobilization. *Journal of cell science*, 125(Pt 15):3568–3577.

Bieth, E., Eddiry, S., Gaston, V., Lorenzini, F., Buffet, A., Conte Auriol, F., Molinas, C., Cailley, D., Rooryck, C., Arveiler, B., Cavaillé, J., Salles, J. P., and Tauber, M. (2015). Highly restricted deletion of the SNORD116 region is implicated in Prader-Willi Syndrome. *European journal of human genetics : EJHG*, 23(2):252–5.

Birse, R. T., Choi, J., Reardon, K., Rodriguez, J., Graham, S., Diop, S., Ocorr, K., Bodmer, R., and Oldham, S. (2010). High-fat-diet-induced obesity and heart dysfunction are regulated by the TOR pathway in Drosophila. *Cell metabolism*, 12(5):533–544.

Blundell, J. E., Campbell, D. B., Lesham, M., and Tozer, R. (1975). Comparison of the time course of the anorectic effect of fenfluramine and amphetamine with drug levels in blood. *The Journal of pharmacy and pharmacology*, 27(3):187–192.

Blundell, J. E. and Latham, C. J. (1979). Serotonergic influences on food intake: effect of 5-hydroxytryptophan on parameters of feeding behaviour in deprived and free-feeding rats. *Pharmacology, biochemistry, and behavior*, 11(4):431–437.

Bogardus, C. (2009). Missing heritability and GWAS utility. *Obesity (Silver Spring, Md.)*, 17(2):209–210.

Bogart, K. and Andrews, J. (2006). *Extraction of Total RNA from Drosophila*, volume 2006. Center for Genomics and Bioinformatics.

Bohni, R., Riesgo-Escovar, J., Oldham, S., Brogiolo, W., Stocker, H., Andruss, B. F., Beckingham, K., and Hafen, E. (1999). Autonomous control of cell and organ size by CHICO, a Drosophila homolog of vertebrate IRS1-4. *Cell*, 97(7):865–875.

Borman, A. D., Pearce, L. R., Mackay, D. S., Nagel-Wolfrum, K., Davidson, A. E., Henderson, R., Garg, S., Waseem, N. H., Webster, A. R., Plagnol, V., Wolfrum, U., Farooqi, I. S., and Moore, A. T. (2014). A homozygous mutation in the TUB gene associated with retinal dystrophy and obesity. *Human mutation*, 35(3):289–93.

Boulay, J.-L., O'Shea, J. J., and Paul, W. E. (2003). Molecular Phylogeny within Type I Cytokines and Their Cognate Receptors. *Immunity*, 19(2):159–163.

Brand, A. H. and Perrimon, N. (1993). Targeted gene expression as a means of altering cell fates and generating dominant phenotypes. *Development (Cambridge, England)*, 118(2):401–15.

Breisch, S. T., Zemlan, F. P., and Hoebel, B. G. (1976). Hyperphagia and obesity following serotonin depletion by intraventricular p-chlorophenylalanine. *Science (New York, N.Y.)*, 192(4237):382–385.

Brighton, P. J., Szekeres, P. G., and Willars, G. B. (2004). Neuromedin U and its receptors: structure, function, and physiological roles. *Pharmacological reviews*, 56(2):231–248.

Brogiolo, W., Stocker, H., Ikeya, T., Rintelen, F., Fernandez, R., and Hafen, E. (2001). An evolutionarily conserved function of the Drosophila insulin receptor and insulin-like peptides in growth control. *Curr Biol*, 11.

Brown, M. R., Crim, J. W., Arata, R. C., Cai, H. N., Chun, C., and Shen, P. (1999). Identification of a Drosophila brain-gut peptide related to the neuropeptide Y family. *Peptides*, 20(9):1035–1042.

Brüning, J. C., Gautam, D., Burks, D. J., Gillette, J., Schubert, M., Orban, P. C., Klein, R., Krone, W., Müller-Wieland, D., and Kahn, C. R. (2000). Role of Brain Insulin Receptor in Control of Body Weight and Reproduction. *Science*, 289(5487):2122 LP – 2125.

Busson, D. and Pret, A.-M. (2007). GAL4/UAS Targeted Gene Expression for Studying Drosophila Hedgehog Signaling BT - Hedgehog Signaling Protocols. In Horabin, J. I., editor, *Hedgehog Signaling Protocols*, pages 161–201. Humana Press, Totowa.

Butler, M., Rafi, S., Hossain, W., Stephan, D., and Manzardo, A. (2015). Whole Exome Sequencing in Females with Autism Implicates Novel and Candidate Genes. *International Journal of Molecular Sciences*, 16(1):1312–1335.

Caravaca, J. M. and Lei, E. P. (2016). Maintenance of a Drosophila melanogaster Population Cage. *Journal of visualized experiments : JoVE*, 109:53756.

Carroll, J. B., Bates, G. P., Steffan, J., Saft, C., and Tabrizi, S. J. (2015). Treating the whole body in Huntington's disease. *The Lancet Neurology*, 14(11):1135–1142.

Carroll, K., Gomez, C., and Shapiro, L. (2004). Tubby proteins: the plot thickens. *Nature reviews. Molecular cell biology*, 5(1):55–63.

Castro, M. G. and Morrison, E. (1997). Post-translational processing of proopiomelanocortin in the pituitary and in the brain. *Critical reviews in neurobiology*, 11(1):35–57.

Chakraborty, R., Vepuri, V., Mhatre, S. D., Paddock, B. E., Miller, S., Michelson, S. J., Delvadia, R., Desai, A., Vinokur, M., Melicharek, D. J., Utreja, S., Khandelwal, P., Ansaloni, S., Goldstein, L. E., Moir, R. D., Lee, J. C., Tabb, L. P., Saunders, A. J., and Marenda, D. R. (2011). Characterization of a Drosophila Alzheimer's Disease Model: Pharmacological Rescue of Cognitive Defects. *PLoS ONE*, 6(6):e20799.

Challis, B. G., Coll, A. P., Yeo, G. S. H., Pinnock, S. B., Dickson, S. L., Thresher, R. R., Dixon, J., Zahn, D., Rochford, J. J., White, A., Oliver, R. L., Millington, G., Aparicio, S. A., Colledge, W. H., Russ, A. P., Carlton, M. B., and O'Rahilly, S. (2004). Mice lacking pro-opiomelanocortin are sensitive to high-fat feeding but respond normally to the acute anorectic effects of peptide-YY(3-36). *Proceedings of the National Academy of Sciences of the United States of America*, 101(13):4695–4700.

Cherian, P. T., Al-Khairi, I., Sriraman, D., Al-Enezi, A., Al-Sultan, D., AlOtaibi, M., Al-Enezi, S., Tuomilehto, J., Al-Mulla, F., Abubaker, J. A., and Abu-Farha, M. (2018). Increased Circulation and Adipose Tissue Levels of DNAJC27/RBJ in Obesity and Type 2-Diabetes. *Frontiers in endocrinology*, 9:423.

Chua, S. C. J., Chung, W. K., Wu-Peng, X. S., Zhang, Y., Liu, S. M., Tartaglia, L., and Leibel, R. L. (1996). Phenotypes of mouse diabetes and rat fatty due to mutations in the OB (leptin) receptor. *Science (New York, N.Y.)*, 271(5251):994–996.

Chung, W. K. (2012). An overview of mongenic and syndromic obesities in humans. *Pediatric blood & cancer*, 58(1):122–8.

Clément, K., Biebermann, H., Farooqi, I. S., Van der Ploeg, L., Wolters, B., Poitou, C., Puder, L., Fiedorek, F., Gottesdiener, K., Kleinau, G., Heyder, N., Scheerer, P., Blume-Peytavi, U., Jahnke, I., Sharma, S., Mokrosinski, J., Wiegand, S., Müller, A., Weiß, K., Mai, K., Spranger, J., Grüters, A., Blankenstein, O., Krude, H., and Kühnen, P. (2018). MC4R agonism promotes durable weight loss in patients with leptin receptor deficiency. *Nature Medicine*, 24(5):551–555.

Clements, J., Hens, K., Francis, C., Schellens, A., and Callaerts, P. (2008). Conserved role for the Drosophila Pax6 homolog Eyeless in differentiation and function of insulin-producing neurons. *Proceedings of the National Academy of Sciences*, 105(42):16183 LP – 16188.

Cone, R. D. (2005). Anatomy and regulation of the central melanocortin system. *Nature neuroscience*, 8(5):571–8.

Conesa, A., Madrigal, P., Tarazona, S., Gomez-Cabrero, D., Cervera, A., McPherson, A., Szcześniak, M. W., Gaffney, D. J., Elo, L. L., Zhang, X., and Mortazavi, A. (2016). A survey of best practices for RNA-seq data analysis. *Genome Biology*, 17(1):13.

Cong, L., Ran, F. A., Cox, D., Lin, S., Barretto, R., and Habib, N. (2013). Multiplex genome engineering using CRISPR/Cas systems. *Science*, 339.

Considine, R. V., Sinha, M. K., Heiman, M. L., Kriauciunas, A., Stephens, T. W., Nyce, M. R., Ohannesian, J. P., Marco, C. C., McKee, L. J., and Bauer, T. L. (1996). Serum immunoreactive-leptin concentrations in normal-weight and obese humans. *The New England journal of medicine*, 334(5):292–295.

Cowley, M. A., Smart, J. L., Rubinstein, M., Cerdan, M. G., Diano, S., Horvath, T. L., Cone, R. D., and Low, M. J. (2001). Leptin activates anorexigenic POMC neurons through a neural network in the arcuate nucleus. *Nature*, 411(6836):480–484.

Crivat, G., Lizunov, V. A., Li, C. R., Stenkula, K. G., Zimmerberg, J., Cushman, S. W., and Pick, L. (2013). Insulin Stimulates Translocation of Human GLUT4 to the Membrane in Fat Bodies of Transgenic Drosophila melanogaster. *PLoS ONE*, 8(11):e77953.

Cvetkovic, V., Poncet, F., Fellmann, D., Griffond, B., and Risold, P. Y. (2003). Diencephalic neurons producing melanin-concentrating hormone are influenced by local and multiple extra-hypothalamic tachykininergic projections through the neurokinin 3 receptor. *Neuroscience*, 119(4):1113–1145.

Di Angelantonio, E., Bhupathiraju, S. N., Wormser, D., Gao, P., Kaptoge, S., de Gonzalez, A. B., Cairns, B. J., Huxley, R., Jackson, C. L., Joshy, G., Lewington, S., Manson, J. E., Murphy, N., Patel, A. V., Samet, J. M., Woodward, M., Zheng, W., Zhou, M., Bansal, N., Barricarte, A., Carter, B., Cerhan, J. R., Collins, R., Smith, G. D., Fang, X., Franco, O. H., Green, J., Halsey, J., Hildebrand, J. S., Jung, K. J., Korda, R. J., McLerran, D. F.,

Moore, S. C., O'Keeffe, L. M., Paige, E., Ramond, A., Reeves, G. K., Rolland, B., Sacerdote, C., Sattar, N., Sofianopoulou, E., Stevens, J., Thun, M., Ueshima, H., Yang, L., Yun, Y. D., Willeit, P., Banks, E., Beral, V., Chen, Z., Gapstur, S. M., Gunter, M. J., Hartge, P., Jee, S. H., Lam, T.-H., Peto, R., Potter, J. D., Willett, W. C., Thompson, S. G., Danesh, J., and Hu, F. B. (2018). Body-mass index and all-cause mortality: individual-participant-data meta-analysis of 239 prospective studies in four continents. *The Lancet*, 388(10046):776–786.

Dickinson, M. E., Flenniken, A. M., Ji, X., Teboul, L., Wong, M. D., White, J. K., Meehan, T. F., Weninger, W. J., Westerberg, H., Adissu, H., Baker, C. N., Bower, L., Brown, J. M., Caddle, L. B., Chiani, F., Clary, D., Cleak, J., Daly, M. J., Denegre, J. M., Doe, B., Dolan, M. E., Edie, S. M., Fuchs, H., Gailus-Durner, V., Galli, A., Gambadoro, A., Gallegos, J., Guo, S., Horner, N. R., Hsu, C.-W., Johnson, S. J., Kalaga, S., Keith, L. C., Lanoue, L., Lawson, T. N., Lek, M., Mark, M., Marschall, S., Mason, J., McElwee, M. L., Newbigging, S., Nutter, L. M. J., Peterson, K. A., Ramirez-Solis, R., Rowland, D. J., Ryder, E., Samocha, K. E., Seavitt, J. R., Selloum, M., Szoke-Kovacs, Z., Tamura, M., Trainor, A. G., Tudose, I., Wakana, S., Warren, J., Wendling, O., West, D. B., Wong, L., Yoshiki, A., MacArthur, D. G., Tocchini-Valentini, G. P., Gao, X., Flicek, P., Bradley, A., Skarnes, W. C., Justice, M. J., Parkinson, H. E., Moore, M., Wells, S., Braun, R. E., Svenson, K. L., de Angelis, M. H., Herault, Y., Mohun, T., Mallon, A.-M., Henkelman, R. M., Brown, S. D. M., Adams, D. J., Lloyd, K. C. K., McKerlie, C., Beaudet, A. L., Bućan, M., Murray, S. A., Henkelman, R. M., Brown, S. D. M., Adams, D. J., Lloyd, K. C. K., McKerlie, C., Beaudet, A. L., Bućan, M., and Murray, S. A. (2016). High-throughput discovery of novel developmental phenotypes. *Nature*, 537(7621):508–514.

Dietzl, G., Chen, D., Schnorrer, F., Su, K.-C., Barinova, Y., Fellner, M., Gasser, B., Kinsey, K., Oppel, S., Scheiblauer, S., Couto, A., Marra, V., Keleman, K., and Dickson, B. J. (2007). A genome-wide transgenic RNAi library for conditional gene inactivation in Drosophila. *Nature*, 448(7150):151–156.

Doche, M. E., Bochukova, E. G., Su, H.-W., Pearce, L. R., Keogh, J. M., Henning, E., Cline, J. M., Saeed, S., Dale, A., Cheetham, T., Barroso, I., Argetsinger, L. S., O'Rahilly, S., Rui, L., Carter-Su, C., and Farooqi, I. S. (2012). Human SH2B1 mutations are associated with maladaptive behaviors and obesity. *The Journal of clinical investigation*, 122(12):4732–6.

Doco-Fenzy, M., Leroy, C., Schneider, A., Petit, F., Delrue, M.-A., Andrieux, J., Perrin-Sabourin, L., Landais, E., Aboura, A., Puechberty, J., Girard, M., Tournaire, M., Sanchez, E., Rooryck, C., Ameil, A., Goossens, M., Jonveaux, P., Lefort, G., Taine, L., Cailley, D., Gaillard, D., Leheup, B., Sarda, P., and Geneviève, D. (2014). Early-onset obesity and paternal 2pter deletion encompassing the ACP1, TMEM18, and MYT1L genes. *European journal of human genetics : EJHG*, 22(4):471–9.

Doench, J. G., Fusi, N., Sullender, M., Hegde, M., Vaimberg, E. W., and Donovan, K. F. (2016). Optimized sgRNA design to maximize activity and minimize off-target effects of CRISPR-Cas9. *Nat Biotechnol*, 34.

Donovan, M. H. and Tecott, L. H. (2013). Serotonin and the regulation of mammalian energy balance. *Frontiers in Neuroscience*, 7.

Duffy, J. B. (2002). GAL4 system in Drosophila: a fly geneticist's Swiss army knife. *Genesis (New York, N.Y. : 2000)*, 34(1-2):1–15.

Echwald, S. M., Rasmussen, S. B., Sorensen, T. I., Andersen, T., Tybjaerg-Hansen, A., Clausen, J. O., Hansen, L., Hansen, T., and Pedersen, O. (1997). Identification of two novel missense mutations in the human OB gene. *International journal of obesity and related metabolic disorders : journal of the International Association for the Study of Obesity*, 21(4):321–326.

Edgecomb, R. S., Harth, C. E., and Schneiderman, A. M. (1994). Regulation of feeding behavior in adult Drosophila melanogaster varies with feeding regime and nutritional state. *The Journal of experimental biology*, 197:215–235.

Ellis, L. L. and Carney, G. E. (2010). Mating alters gene expression patterns in Drosophila melanogaster male heads. *BMC Genomics*, 11(1):1–14.

Ernst, M. B., Wunderlich, C. M., Hess, S., Paehler, M., Mesaros, A., Koralov, S. B., Kleinridders, A., Husch, A., Munzberg, H., Hampel, B., Alber, J., Kloppenburg, P., Bruning, J. C., and Wunderlich, F. T. (2009). Enhanced Stat3 Activation in POMC Neurons Provokes Negative Feedback Inhibition of Leptin and InsulinSignaling in Obesity. *Journal of Neuroscience*, 29(37):11582–11593.

Escayg, A., De Waard, M., Lee, D. D., Bichet, D., Wolf, P., Mayer, T., Johnston, J., Baloh, R., Sander, T., and Meisler, M. H. (2000). Coding and Noncoding Variation of the Human Calcium-Channel β4-Subunit Gene CACNB4 in Patients with Idiopathic Generalized Epilepsy and Episodic Ataxia. *The American Journal of Human Genetics*, 66(5):1531–1539.

Farooqi, I. S., Jebb, S. A., Langmack, G., Lawrence, E., Cheetham, C. H., Prentice, A. M., Hughes, I. A., McCamish, M. A., and O'Rahilly, S. (1999). Effects of recombinant leptin therapy in a child with congenital leptin deficiency. *The New England journal of medicine*, 341(12):879–884.

Farooqi, I. S. and O'Rahilly, S. (2014). 20 YEARS OF LEPTIN: Human disorders of leptin action. *Journal of Endocrinology*, 223(1):T63–T70.

Fauny, J. D., Silber, J., and Zider, A. (2005). Drosophila Lipid Storage Droplet 2 gene (Lsd-2) is expressed and controls lipid storage in wing imaginal discs. *Developmental dynamics : an official publication of the American Association of Anatomists*, 232(3):725–732.

Fei, H., Okano, H. J., Li, C., Lee, G. H., Zhao, C., Darnell, R., and Friedman, J. M. (1997). Anatomic localization of alternatively spliced leptin receptors (Ob-R) in mouse brain and other tissues. *Proceedings of the National Academy of Sciences of the United States of America*, 94(13):7001–5.

Feinleib, M., Garrison, R. J., Fabsitz, R., Christian, J. C., Hrubec, Z., Borhani, N. O., Kannel, W. B., Rosenman, R., Schwartz, J. T., and Wagner, J. O. (1977). The NHLBI twin study of cardiovascular disease risk factors: methodology and summary of results. *American journal of epidemiology*, 106(4):284–285.

Ferreira, M. A., Vonk, J. M., Baurecht, H., Marenholz, I., Tian, C., Hoffman, J. D., Helmer, Q., Tillander, A., Ullemar, V., van Dongen, J., Lu, Y., Rüschendorf, F., Esparza-Gordillo, J., Medway, C. W., Mountjoy, E., Burrows, K., Hummel, O., Grosche, S., Brumpton, B. M., Witte, J. S., Hottenga, J.-J., Willemsen, G., Zheng, J., Rodríguez, E., Hotze, M., Franke,

A., Revez, J. A., Beesley, J., Matheson, M. C., Dharmage, S. C., Bain, L. M., Fritsche, L. G., Gabrielsen, M. E., Balliu, B., Nielsen, J. B., Zhou, W., Hveem, K., Langhammer, A., Holmen, O. L., Løset, M., Abecasis, G. R., Willer, C. J., Arnold, A., Homuth, G., Schmidt, C. O., Thompson, P. J., Martin, N. G., Duffy, D. L., Novak, N., Schulz, H., Karrasch, S., Gieger, C., Strauch, K., Melles, R. B., Hinds, D. A., Hübner, N., Weidinger, S., Magnusson, P. K. E., Jansen, R., Jorgenson, E., Lee, Y.-A., Boomsma, D. I., Almqvist, C., Karlsson, R., Koppelman, G. H., Paternoster, L., Almqvist, C., Karlsson, R., Koppelman, G. H., and Paternoster, L. (2017). Shared genetic origin of asthma, hay fever and eczema elucidates allergic disease biology. *Nature Genetics*, 49(12):1752–1757.

Fortini, M. E., Skupski, M. P., Boguski, M. S., and Hariharan, I. K. (2000). A survey of human disease gene counterparts in the Drosophila genome. *The Journal of cell biology*, 150(2):F23–30.

Frayling, T. M., Timpson, N. J., Weedon, M. N., Zeggini, E., Freathy, R. M., Lindgren, C. M., Perry, J. R. B., Elliott, K. S., Lango, H., Rayner, N. W., Shields, B., Harries, L. W., Barrett, J. C., Ellard, S., Groves, C. J., Knight, B., Patch, A.-M., Ness, A. R., Ebrahim, S., Lawlor, D. A., Ring, S. M., Ben-Shlomo, Y., Jarvelin, M.-R., Sovio, U., Bennett, A. J., Melzer, D., Ferrucci, L., Loos, R. J. F., Barroso, I., Wareham, N. J., Karpe, F., Owen, K. R., Cardon, L. R., Walker, M., Hitman, G. A., Palmer, C. N. A., Doney, A. S. F., Morris, A. D., Smith, G. D., Hattersley, A. T., and McCarthy, M. I. (2007). A common variant in the FTO gene is associated with body mass index and predisposes to childhood and adult obesity. *Science*, 316(5826):889–894.

Frederich, R. C., Hamann, A., Anderson, S., Löllmann, B., Lowell, B. B., and Flier, J. S. (1995). Leptin levels reflect body lipid content in mice: evidence for diet-induced resistance to leptin action. *Nature medicine*, 1(12):1311–4.

French, S. A., Story, M., and Jeffery, R. W. (2001). Environmental influences on eating and physical activity. *Annual review of public health*, 22:309–335.

Fridell, Y.-W. C., Hoh, M., Kréneisz, O., Hosier, S., Chang, C., Scantling, D., Mulkey, D. K., and Helfand, S. L. (2009). Increased uncoupling protein (UCP) activity in Drosophila insulin-producing neurons attenuates insulin signaling and extends lifespan. *Aging*, 1(8):699–713.

Fu, Y., Foden, J. A., Khayter, C., Maeder, M. L., Reyon, D., Joung, J. K., and Sander, J. D. (2013). High-frequency off-target mutagenesis induced by CRISPR-Cas nucleases in human cells. *Nature Biotechnology*, 31(9):822–826.

Gadde, K. M., Martin, C. K., Berthoud, H.-R., and Heymsfield, S. B. (2018). Obesity: Pathophysiology and Management. *Journal of the American College of Cardiology*, 71(1):69–84.

Gandotra, S., Le Dour, C., Bottomley, W., Cervera, P., Giral, P., Reznik, Y., Charpentier, G., Auclair, M., Delépine, M., Barroso, I., Semple, R. K., Lathrop, M., Lascols, O., Capeau, J., O'Rahilly, S., Magré, J., Savage, D. B., and Vigouroux, C. (2011). Perilipin Deficiency and Autosomal Dominant Partial Lipodystrophy. *New England Journal of Medicine*, 364(8):740–748.

Gao, X., Schwarzschild, M. A., Wang, H., and Ascherio, A. (2009). Obesity and restless legs syndrome in men and women. *Neurology*, 72(14):1255–1261.

Garczynski, S. F., Brown, M. R., Shen, P., Murray, T. F., and Crim, J. W. (2002). Characterization of a functional neuropeptide F receptor from Drosophila melanogaster. *Peptides*, 23(4):773–780.

Gasque, G., Conway, S., Huang, J., Rao, Y., and Vosshall, L. B. (2013). Small molecule drug screening in Drosophila identifies the 5HT2A receptor as a feeding modulation target. *Scientific reports*, 3:srep02120.

GBD (2017). Health Effects of Overweight and Obesity in 195 Countries over 25 Years. *New England Journal of Medicine*, 377(1):13–27.

Ge, D. T., Tipping, C., Brodsky, M. H., and Zamore, P. D. (2016). Rapid Screening for CRISPR-Directed Editing of the Drosophila Genome Using white Co-conversion. *G3 (Bethesda, Md.)*, 6(10):3197–3206.

Gelling, R. W., Du, X. Q., Dichmann, D. S., Rømer, J., Huang, H., Cui, L., Obici, S., Tang, B., Holst, J. J., Fledelius, C., Johansen, P. B., Rossetti, L., Jelicks, L. A., Serup, P., Nishimura, E., and Charron, M. J. (2003). Lower blood glucose, hyperglucagonemia, and pancreatic α cell hyperplasia in glucagon receptor knockout mice. *Proceedings of the National Academy of Sciences*, 100(3):1438 LP – 1443.

Geyer, M. A., Puerto, A., Menkes, D. B., Segal, D. S., and Mandell, A. J. (1976). Behavioral studies following lesions of the mesolimbic and mesostriatal serotonergic pathways. *Brain research*, 106(2):257–269.

Graham, M., Shutter, J. R., Sarmiento, U., Sarosi, I., and Stark, K. L. (1997). Overexpression of Agrt leads to obesity in transgenic mice. *Nature genetics*, 17(3):273–274.

Graham, P. and Pick, L. (2017). Drosophila as a Model for Diabetes and Diseases of Insulin Resistance. *Current topics in developmental biology*, 121:397–419.

Gramates, L. S., Marygold, S. J., Santos, G. D., Urbano, J.-M., Antonazzo, G., Matthews, B. B., Rey, A. J., Tabone, C. J., Crosby, M. A., Emmert, D. B., Falls, K., Goodman, J. L., Hu, Y., Ponting, L., Schroeder, A. J., Strelets, V. B., Thurmond, J., Zhou, P., and the FlyBase Consortium (2017). FlyBase at 25: looking to the future. *Nucleic acids research*, 45(D1):D663–D671.

Gratz, S. J., Cummings, A. M., Nguyen, J. N., Hamm, D. C., Donohue, L. K., Harrison, M. M., Wildonger, J., and O'Connor-Giles, K. M. (2013). Genome Engineering of Drosophila with the CRISPR RNA-Guided Cas9 Nuclease. *Genetics*, 194(4):1029–1035.

Graveley, B. R., May, G., Brooks, A. N., Carlson, J. W., Cherbas, L., Davis, C. A., Duff, M., Eads, B., Landolin, J., Sandler, J., Wan, K. H., Andrews, J., Brenner, S. E., Cherbas, P., Gingeras, T. R., Hoskins, R., Kaufman, T., and Celniker, S. E. (2011). The D. melanogaster transcriptome: modENCODE RNA-Seq data for dissected tissues.

Gray, J., Yeo, G., Hung, C., Keogh, J., Clayton, P., Banerjee, K., McAulay, A., O'Rahilly, S., and Farooqi, I. S. (2007). Functional characterization of human NTRK2 mutations identified in patients with severe early-onset obesity. *International Journal of Obesity*, 31(2):359–364.

Green, E. W., Fedele, G., Giorgini, F., and Kyriacou, C. P. (2014). A Drosophila RNAi collection is subject to dominant phenotypic effects. *Nature Methods*, 11(3):222–223.

Grinspoon, S., Thomas, E., Pitts, S., Gross, E., Mickley, D., Miller, K., Herzog, D., and Klibanski, A. (2000). Prevalence and predictive factors for regional osteopenia in women with anorexia nervosa. *Annals of internal medicine*, 133(10):790–4.

Gronke, S., Beller, M., Fellert, S., Ramakrishnan, H., Jackle, H., and Kuhnlein, R. P. (2003). Control of fat storage by a Drosophila PAT domain protein. *Current biology : CB*, 13(7):603–606.

Gronke, S., Mildner, A., Fellert, S., Tennagels, N., Petry, S., Muller, G., Jackle, H., and Kuhnlein, R. P. (2005). Brummer lipase is an evolutionary conserved fat storage regulator in Drosophila. *Cell metabolism*, 1(5):323–330.

Gronke, S., Muller, G., Hirsch, J., Fellert, S., Andreou, A., Haase, T., Jackle, H., and Kuhnlein, R. P. (2007). Dual lipolytic control of body fat storage and mobilization in Drosophila. *PLoS biology*, 5(6):e137.

Groth, A. C., Fish, M., Nusse, R., and Calos, M. P. (2004). Construction of transgenic Drosophila by using the site-specific integrase from phage phiC31. *Genetics*, 166(4):1775–1782.

Haemmerle, G., Lass, A., Zimmermann, R., Gorkiewicz, G., Meyer, C., Rozman, J., Heldmaier, G., Maier, R., Theussl, C., Eder, S., Kratky, D., Wagner, E. F., Klingenspor, M., Hoefler, G., and Zechner, R. (2006). Defective lipolysis and altered energy metabolism in mice lacking adipose triglyceride lipase. *Science (New York, N.Y.)*, 312(5774):734–737.

Haeussler, M., Schönig, K., Eckert, H., Eschstruth, A., Mianné, J., Renaud, J.-B., Schneider-Maunoury, S., Shkumatava, A., Teboul, L., Kent, J., Joly, J.-S., and Concordet, J.-P. (2016). Evaluation of off-target and on-target scoring algorithms and integration into the guide RNA selection tool CRISPOR. *Genome Biology*, 17(1):148.

Hahn, T. M., Breininger, J. F., Baskin, D. G., and Schwartz, M. W. (1998). Coexpression of Agrp and NPY in fasting-activated hypothalamic neurons. *Nature Neuroscience*, 1(4):271–272.

Halaas, J. L., Gajiwala, K. S., Maffei, M., Cohen, S. L., Chait, B. T., Rabinowitz, D., Lallone, R. L., Burley, S. K., and Friedman, J. M. (1995). Weight-reducing effects of the plasma protein encoded by the obese gene. *Science (New York, N.Y.)*, 269(5223):543–546.

Hamdan, F. F., Srour, M., Capo-Chichi, J.-M., Daoud, H., Nassif, C., Patry, L., Massicotte, C., Ambalavanan, A., Spiegelman, D., Diallo, O., Henrion, E., Dionne-Laporte, A., Fougerat, A., Pshezhetsky, A. V., Venkateswaran, S., Rouleau, G. A., and Michaud, J. L. (2014). De Novo Mutations in Moderate or Severe Intellectual Disability. *PLoS Genetics*, 10(10):e1004772.

Hanada, R., Teranishi, H., Pearson, J. T., Kurokawa, M., Hosoda, H., Fukushima, N., Fukue, Y., Serino, R., Fujihara, H., Ueta, Y., Ikawa, M., Okabe, M., Murakami, N., Shirai, M., Yoshimatsu, H., Kangawa, K., and Kojima, M. (2004). Neuromedin U has a novel anorexigenic effect independent of the leptin signaling pathway. *Nature medicine*, 10(10):1067–1073.

Harrison, D. A., McCoon, P. E., Binari, R., Gilman, M., and Perrimon, N. (1998). Drosophila unpaired encodes a secreted protein that activates the JAK signaling pathway. *Genes & development*, 12(20):3252–3263.

Heigwer, F., Port, F., and Boutros, M. (2018). RNA Interference (RNAi) Screening in Drosophila. *Genetics*, 208(3):853 LP – 874.

Heinrichsen, E. T. and Haddad, G. G. (2012). Role of High-Fat Diet in Stress Response of Drosophila. *PLoS ONE*, 7(8):e42587.

Held, L. (2017). *Deep Homology?* Cambridge University Press, 1 edition.

Henry, F. E., Sugino, K., Tozer, A., Branco, T., and Sternson, S. M. (2015). Cell type-specific transcriptomics of hypothalamic energy-sensing neuron responses to weight-loss. *eLife*, 4.

Henry, M., Ghibaudi, L., Gao, J., and Hwa, J. J. (2005). Energy metabolic profile of mice after chronic activation of central NPY Y1, Y2, or Y5 receptors. *Obesity research*, 13(1):36–47.

Hentges, K. E. and Justice, M. J. (2004). Checks and balancers: balancer chromosomes to facilitate genome annotation. *Trends in genetics*, 20(6):252–9.

Hetherington, A. W. and Ranson, S. W. (1942). The relation of various hypothalamic lesions to adiposity in the rat. *The Journal of Comparative Neurology*, 76(3):475–499.

Heymsfield, S. B., Greenberg, A. S., Fujioka, K., Dixon, R. M., Kushner, R., Hunt, T., Lubina, J. A., Patane, J., Self, B., Hunt, P., and McCamish, M. (1999). Recombinant Leptin for Weight Loss in Obese and Lean Adults. *JAMA*, 282(16):1568.

Hill, A. P., Zuckerman, K. E., and Fombonne, E. (2015). Obesity and Autism. *PEDIATRICS*, 136(6):1051–1061.

Holm, T., Maier, A., Wicks, P., Lang, D., Linke, P., Münch, C., Steinfurth, L., Meyer, R., and Meyer, T. (2013). Severe loss of appetite in amyotrophic lateral sclerosis patients: online self-assessment study. *Interactive journal of medical research*, 2(1):e8.

Howard, A. D., Wang, R., Pong, S. S., Mellin, T. N., Strack, A., Guan, X. M., Zeng, Z., Williams, D. L. J., Feighner, S. D., Nunes, C. N., Murphy, B., Stair, J. N., Yu, H., Jiang, Q., Clements, M. K., Tan, C. P., McKee, K. K., Hreniuk, D. L., McDonald, T. P., Lynch, K. R., Evans, J. F., Austin, C. P., Caskey, C. T., Van der Ploeg, L. H., and Liu, Q. (2000). Identification of receptors for neuromedin U and its role in feeding. *Nature*, 406(6791):70–74.

Hruby, A. and Hu, F. B. (2015). The Epidemiology of Obesity: A Big Picture. *PharmacoEconomics*, 33(7):673–89.

Hsu, P. D., Scott, D. A., Weinstein, J. A., Ran, F. A., Konermann, S., and Agarwala, V. (2013). DNA targeting specificity of RNA-guided Cas9 nucleases. *Nat Biotechnol*, 31.

Hu, X., Zhang, J., Jin, C., Mi, W., Wang, F., Ma, W., Ma, C., Yang, Y., Li, W., Zhang, H., Du, B., Li, K., Liu, C., Wang, L., Lu, T., Zhang, H., Lv, L., Zhang, D., and Yue, W. (2013). Association study of NRXN3 polymorphisms with schizophrenia and risperidone-induced bodyweight gain in Chinese Han population. *Progress in Neuro-Psychopharmacology and Biological Psychiatry*, 43:197–202.

Huang, S. and Czech, M. P. (2007). The GLUT4 Glucose Transporter. *Cell Metabolism*, 5(4):237–252.

Hummel, T. and Klämbt, C. (2008). P-Element Mutagenesis. In *Drosophila Methods and Protocols*, pages 97–117. Humana Press.

Huo, Y.-X., Huang, L., Zhang, D.-F., Yao, Y.-G., Fang, Y.-R., Zhang, C., and Luo, X.-J. (2016). Identification of SLC25A37 as a major depressive disorder risk gene. *Journal of Psychiatric Research*, 83:168–175.

Huvenne, H., Dubern, B., Clément, K., and Poitou, C. (2016). Rare Genetic Forms of Obesity: Clinical Approach and Current Treatments in 2016. *Obesity facts*, 9(3):158–73.

Iliadi, K. G. and Boulianne, G. L. (2010). Age-related behavioral changes in Drosophila. *Annals of the New York Academy of Sciences*, 1197:9–18.

Itskov, P. M. and Ribeiro, C. (2013). The dilemmas of the gourmet fly: the molecular and neuronal mechanisms of feeding and nutrient decision making in Drosophila. *Frontiers in neuroscience*, 7:12.

Ja, W. W. (2018). Satiation and the Sleep-feeding Axis. In *Neurofly*.

Ja, W. W., Carvalho, G. B., Mak, E. M., de la Rosa, N. N., Fang, A. Y., Liong, J. C., Brummel, T., and Benzer, S. (2007). Prandiology of Drosophila and the CAFE assay. *Proceedings of the National Academy of Sciences of the United States of America*, 104(20):8253–8256.

Jackson, R. S., Creemers, J. W., Ohagi, S., Raffin-Sanson, M. L., Sanders, L., Montague, C. T., Hutton, J. C., and O'Rahilly, S. (1997). Obesity and impaired prohormone processing associated with mutations in the human prohormone convertase 1 gene. *Nature genetics*, 16(3):303–306.

Jackson Laboratory (2018). Mouse Genome Informatics Web Site.

Jehrke, L., Stewart, F. A., Droste, A., and Beller, M. (2018). The impact of genome variation and diet on the metabolic phenotype and microbiome composition of Drosophila melanogaster. *Scientific Reports*, 8(1):6215.

Jennings, B. (2011). Drosophila – a versatile model in biology and medicine. *Materials Today*, 14(5):190–195.

Jiang, G. and Zhang, B. B. (2003). Glucagon and regulation of glucose metabolism. *American journal of physiology. Endocrinology and metabolism*, 284(4):E671–8.

Jiang, Y., Wilk, J. B., Borecki, I., Williamson, S., DeStefano, A. L., Xu, G., Liu, J., Ellison, R. C., Province, M., and Myers, R. H. (2004). Common variants in the 5' region of the leptin gene are associated with body mass index in men from the National Heart, Lung, and Blood Institute Family Heart Study. *American journal of human genetics*, 75(2):220–230.

Johnson, E. O., Hancock, D. B., Gaddis, N. C., Levy, J. L., Page, G., Novak, S. P., Glasheen, C., Saccone, N. L., Rice, J. P., Moreau, M. P., Doheny, K. F., Romm, J. M., Brooks, A. I., Aouizerat, B. E., Bierut, L. J., and Kral, A. H. (2015). Novel Genetic Locus Implicated for HIV-1 Acquisition with Putative Regulatory Links to HIV Replication and Infectivity: A Genome-Wide Association Study. *PLOS ONE*, 10(3):e0118149.

Jönsson, E. G., Edman-Ahlbom, B., Sillén, A., Gunnar, A., Kulle, B., Frigessi, A., Vares, M., Ekholm, B., Wode-Helgodt, B., Schumacher, J., Cichon, S., Agartz, I., Sedvall, G. C., Hall, H., and Terenius, L. (2006). Brain-derived neurotrophic factor gene (BDNF) variants and schizophrenia: An association study. *Progress in Neuro-Psychopharmacology and Biological Psychiatry*, 30(5):924–933.

Jonsson, L., Zettergren, A., Pettersson, E., Hovey, D., Anckarsäter, H., Westberg, L., Lichtenstein, P., Lundström, S., and Melke, J. (2014). Association study between autistic-like traits and polymorphisms in the autism candidate regions RELN, CNTNAP2, SHANK3, and CDH9/10. *Molecular Autism*, 5(1):55.

Jumbo-Lucioni, P., Ayroles, J. F., Chambers, M. M., Jordan, K. W., Leips, J., Mackay, T. F., and De Luca, M. (2010). Systems genetics analysis of body weight and energy metabolism traits in Drosophila melanogaster. *BMC genomics*, 11:297.

Kadam, P. and Bhalerao, S. (2010). Sample size calculation. *International journal of Ayurveda research*, 1(1):55–7.

Kashihara, K. (2006). Weight loss in Parkinson's disease. *Journal of neurology*, 253 Suppl(S7):VII38–41.

Kassis, J. A. (1990). Spatial and temporal control elements of the Drosophila engrailed gene. *Genes & development*, 4(3):433–43.

Kelleher, R. J., III, and Shen, J. (2017). Presenilin-1 mutations and Alzheimer's disease. *Proceedings of the National Academy of Sciences of the United States of America*, 114(4):629.

Kernie, S. G., Liebl, D. J., and Parada, L. F. (2000). BDNF regulates eating behavior and locomotor activity in mice. *The EMBO Journal*, 19(6):1290–1300.

King-Jones, K. and Huynh, N. (2018). A Drosophila CRISPR/Cas9 toolkit for conditionally manipulating gene expression in the prothoracic gland as a test case for polytene tissues. *bioRxiv*.

Klitenick, M. A. and Wirtshafter, D. (1988). Comparative studies of the ingestive behaviors produced by microinjections of muscimol into the midbrain raphe nuclei of the ventral tegmental area of the rat. *Life sciences*, 42(7):775–782.

Kohno, D., Gao, H.-Z., Muroya, S., Kikuyama, S., and Yada, T. (2003). Ghrelin directly interacts with neuropeptide-Y-containing neurons in the rat arcuate nucleus: Ca2+ signaling via protein kinase A and N-type channel-dependent mechanisms and cross-talk with leptin and orexin. *Diabetes*, 52(4):948–56.

Kowalski, T. J., Spar, B. D., Markowitz, L., Maguire, M., Golovko, A., Yang, S., Farley, C., Cook, J. A., Tetzloff, G., Hoos, L., Del Vecchio, R. A., Kazdoba, T. M., McCool, M. F., Hwa, J. J., Hyde, L. A., Davis, H., Vassileva, G., Hedrick, J. A., and Gustafson, E. L. (2005). Transgenic overexpression of neuromedin U promotes leanness and hypophagia in mice. *The Journal of endocrinology*, 185(1):151–64.

Kreneisz, O., Chen, X., Fridell, Y.-W. C., and Mulkey, D. K. (2010). Glucose increases activity and Ca2+ in insulin-producing cells of adult Drosophila. *NeuroReport*, 21(17):1116–1120.

Kristensen, T. N. (2005). Genome-Wide Analysis on Inbreeding Effects on Gene Expression in Drosophila melanogaster. *Genetics*, 171(1):157–167.

Krude, H., Biebermann, H., Luck, W., Horn, R., Brabant, G., and Gruters, A. (1998). Severe early-onset obesity, adrenal insufficiency and red hair pigmentation caused by POMC mutations in humans. *Nature genetics*, 19(2):155–157.

Kubli, E. (2010). Sexual Behavior: Dietary Food Switch Induced by Sex. *Current Biology*, 20:R474–R476.

Kühnlein, R. P. (2012). Lipid droplet-based storage fat metabolism in Drosophila: Thematic Review Series: Lipid Droplet Synthesis and Metabolism: from Yeast to Man. *Journal of Lipid Research*, 53(8):1430–1436.

Larder, R., Sim, M. F. M., Gulati, P., Antrobus, R., Tung, Y. C. L., Rimmington, D., Ayuso, E., Polex-Wolf, J., Lam, B. Y. H., Dias, C., Logan, D. W., Virtue, S., Bosch, F., Yeo, G. S. H., Saudek, V., O'Rahilly, S., and Coll, A. P. (2017). Obesity-associated gene TMEM18 has a role in the central control of appetite and body weight regulation. *Proceedings of the National Academy of Sciences*, 114(35):9421–9426.

Lazareva, A. A., Roman, G., Mattox, W., Hardin, P. E., and Dauwalder, B. (2007). A role for the adult fat body in Drosophila male courtship behavior. *PLoS genetics*, 3(1):e16.

Lee, A. W. S., Hengstler, H., Schwald, K., Berriel-Diaz, M., Loreth, D., Kirsch, M., Kretz, O., Haas, C. A., de Angelis, M. H., Herzig, S., Brümmendorf, T., Klingenspor, M., Rathjen, F. G., Rozman, J., Nicholson, G., Cox, R. D., and Schäfer, M. K. E. (2012). Functional Inactivation of the Genome-Wide Association Study Obesity Gene Neuronal Growth Regulator 1 in Mice Causes a Body Mass Phenotype. *PLoS ONE*, 7(7):e41537.

Lee, G. and Park, J. H. (2004). Hemolymph Sugar Homeostasis and Starvation-Induced Hyperactivity Affected by Genetic Manipulations of the Adipokinetic Hormone-Encoding Gene in Drosophila melanogaster. *Genetics*, 167(1):311 LP – 323.

Lee, K.-S., You, K.-H., Choo, J.-K., Han, Y.-M., and Yu, K. (2004). Drosophila short neuropeptide F regulates food intake and body size. *The Journal of biological chemistry*, 279(49):50781–50789.

Li, J., Houseknecht, K. L., Stenbit, A. E., Katz, E. B., and Charron, M. J. (2000). Reduced glucose uptake precedes insulin signaling defects in adipocytes from heterozygous GLUT4 knockout mice. *FASEB journal : official publication of the Federation of American Societies for Experimental Biology*, 14(9):1117–25.

Li, J. J., Ferry, R. J., Diao, S., Xue, B., Bahouth, S. W., and Liao, F.-F. (2015). Nedd4 Haploinsufficient Mice Display Moderate Insulin Resistance, Enhanced Lipolysis, and Protection Against High-Fat Diet-Induced Obesity. *Endocrinology*, 156(4):1283–1291.

Li, W. D., Reed, D. R., Lee, J. H., Xu, W., Kilker, R. L., Sodam, B. R., and Price, R. A. (1999). Sequence variants in the 5' flanking region of the leptin gene are associated with obesity in women. *Annals of human genetics*, 63(Pt 3):227–234.

Li, X., Tamashiro, K. L. K., Liu, Z., Bello, N. T., Wang, X., Aja, S., Bi, S., Ladenheim, E. E., Ross, C. A., Moran, T. H., and Smith, W. W. (2012). A novel obesity model: synphilin-1-induced hyperphagia and obesity in mice. *International Journal of Obesity*, 36(9):1215–1221.

Lill, C. M. (2016). Genetics of Parkinson's disease. *Molecular and Cellular Probes*, 30(6):386–396.

Liu, J., Li, T., Yang, D., Ma, R., Moran, T. H., and Smith, W. W. (2012). Synphilin-1 alters metabolic homeostasis in a novel Drosophila obesity model. *International Journal of Obesity*, 36(12):1529–1536.

Liu, Y., Luo, J., Carlsson, M. A., and Nässel, D. R. (2015). Serotonin and insulin-like peptides modulate leucokinin-producing neurons that affect feeding and water homeostasis in Drosophila. *Journal of Comparative Neurology*, 523(12):1840–1863.

Locke, A. E., Kahali, B., Berndt, S. I., Justice, A. E., Pers, T. H., Day, F. R., Powell, C., Vedantam, S., Buchkovich, M. L., Yang, J., Croteau-Chonka, D. C., Esko, T., Fall, T., Ferreira, T., Gustafsson, S., Kutalik, Z., Luan, J., Mägi, R., Randall, J. C., Winkler, T. W., Wood, A. R., Workalemahu, T., Faul, J. D., Smith, J. A., Hua Zhao, J., Zhao, W., Chen, J., Fehrmann, R., Hedman, Å. K., Karjalainen, J., Schmidt, E. M., Absher, D., Amin, N., Anderson, D., Beekman, M., Bolton, J. L., Bragg-Gresham, J. L., Buyske, S., Demirkan, A., Deng, G., Ehret, G. B., Feenstra, B., Feitosa, M. F., Fischer, K., Goel, A., Gong, J., Jackson, A. U., Kanoni, S., Kleber, M. E., Kristiansson, K., Lim, U., Lotay, V., Mangino, M., Mateo Leach, I., Medina-Gomez, C., Medland, S. E., Nalls, M. A., Palmer, C. D., Pasko, D., Pechlivanis, S., Peters, M. J., Prokopenko, I., Shungin, D., Stančáková, A., Strawbridge, R. J., Ju Sung, Y., Tanaka, T., Teumer, A., Trompet, S., van der Laan, S. W., van Setten, J., Van Vliet-Ostaptchouk, J. V., Wang, Z., Yengo, L., Zhang, W., Isaacs, A., Albrecht, E., Ärnlöv, J., Arscott, G. M., Attwood, A. P., Bandinelli, S., Barrett, A., Bas, I. N., Bellis, C., Bennett, A. J., Berne, C., Blagieva, R., Blüher, M., Böhringer, S., Bonnycastle, L. L., Böttcher, Y., Boyd, H. A., Bruinenberg, M., Caspersen, I. H., Ida Chen, Y.-D., Clarke, R., Daw, E. W., de Craen, A. J. M., Delgado, G., Dimitriou, M., Doney, A. S. F., Eklund, N., Estrada, K., Eury, E., Folkersen, L., Fraser, R. M., Garcia, M. E., Geller, F., Giedraitis, V., Gigante, B., Go, A. S., Golay, A., Goodall, A. H., Gordon, S. D., Gorski, M., Grabe, H.-J., Grallert, H., Grammer, T. B., Gräßler, J., Grönberg, H., Groves, C. J., Gusto, G., Haessler, J., Hall, P., Haller, T., Hallmans, G., Hartman, C. A., Hassinen, M., Hayward, C., Heard-Costa, N. L., Helmer, Q., Hengstenberg, C., Holmen,

O., Hottenga, J.-J., James, A. L., Jeff, J. M., Johansson, Å., Jolley, J., Juliusdottir, T., Kinnunen, L., Koenig, W., Koskenvuo, M., Kratzer, W., Laitinen, J., Lamina, C., Leander, K., Lee, N. R., Lichtner, P., Lind, L., Lindström, J., Sin Lo, K., Lobbens, S., Lorbeer, R., Lu, Y., Mach, F., Magnusson, P. K. E., Mahajan, A., McArdle, W. L., McLachlan, S., Menni, C., Merger, S., Mihailov, E., Milani, L., Moayyeri, A., Monda, K. L., Morken, M. A., Mulas, A., Müller, G., Müller-Nurasyid, M., Musk, A. W., Nagaraja, R., Nöthen, M. M., Nolte, I. M., Pilz, S., Rayner, N. W., Renstrom, F., Rettig, R., Ried, J. S., Ripke, S., Robertson, N. R., Rose, L. M., Sanna, S., Scharnagl, H., Scholtens, S., Schumacher, F. R., Scott, W. R., Seufferlein, T., Shi, J., Vernon Smith, A., Smolonska, J., Stanton, A. V., Steinthorsdottir, V., Stirrups, K., Stringham, H. M., Sundström, J., Swertz, M. A., Swift, A. J., Syvänen, A.-C., Tan, S.-T., Tayo, B. O., Thorand, B., Thorleifsson, G., Tyrer, J. P., Uh, H.-W., Vandenput, L., Verhulst, F. C., Vermeulen, S. H., Verweij, N., Vonk, J. M., Waite, L. L., Warren, H. R., Waterworth, D., Weedon, M. N., Wilkens, L. R., Willenborg, C., Wilsgaard, T., Wojczynski, M. K., Wong, A., Wright, A. F., Zhang, Q., LifeLines Cohort Study, Brennan, E. P., Choi, M., Dastani, Z., Drong, A. W., Eriksson, P., Franco-Cereceda, A., Gådin, J. R., Gharavi, A. G., Goddard, M. E., Handsaker, R. E., Huang, J., Karpe, F., Kathiresan, S., Keildson, S., Kiryluk, K., Kubo, M., Lee, J.-Y., Liang, L., Lifton, R. P., Ma, B., McCarroll, S. A., McKnight, A. J., Min, J. L., Moffatt, M. F., Montgomery, G. W., Murabito, J. M., Nicholson, G., Nyholt, D. R., Okada, Y., Perry, J. R. B., Dorajoo, R., Reinmaa, E., Salem, R. M., Sandholm, N., Scott, R. A., Stolk, L., Takahashi, A., Tanaka, T., Van't Hooft, F. M., Vinkhuyzen, A. A. E., Westra, H.-J., Zheng, W., Zondervan, K. T., ADIPOGen Consortium, AGEN-BMI Working Group, CARDIOGRAMplusC4D Consortium, CKDGen Consortium, GLGC, ICBP, MAGIC Investigators, MuTHER Consortium, MIGen Consortium, PAGE Consortium, ReproGen Consortium, GENIE Consortium, International Endogene Consortium, Heath, A. C., Arveiler, D., Bakker, S. J. L., Beilby, J., Bergman, R. N., Blangero, J., Bovet, P., Campbell, H., Caulfield, M. J., Cesana, G., Chakravarti, A., Chasman, D. I., Chines, P. S., Collins, F. S., Crawford, D. C., Cupples, L. A., Cusi, D., Danesh, J., de Faire, U., den Ruijter, H. M., Dominiczak, A. F., Erbel, R., Erdmann, J., Eriksson, J. G., Farrall, M., Felix, S. B., Ferrannini, E., Ferrières, J., Ford, I., Forouhi, N. G., Forrester, T., Franco, O. H., Gansevoort, R. T., Gejman, P. V., Gieger, C., Gottesman, O., Gudnason, V., Gyllensten, U., Hall, A. S., Harris, T. B., Hattersley, A. T., Hicks, A. A., Hindorff, L. A., Hingorani, A. D., Hofman, A., Homuth, G., Hovingh, G. K., Humphries, S. E., Hunt, S. C., Hyppönen, E., Illig, T., Jacobs, K. B., Jarvelin, M.-R., Jöckel, K.-H., Johansen, B., Jousilahti, P., Jukema, J. W., Jula, A. M., Kaprio, J., Kastelein, J. J. P., Keinanen-Kiukaanniemi, S. M., Kiemeney, L. A., Knekt, P., Kooner, J. S., Kooperberg, C., Kovacs, P., Kraja, A. T., Kumari, M., Kuusisto, J., Lakka, T. A., Langenberg, C., Le Marchand, L., Lehtimäki, T., Lyssenko, V., Männistö, S., Marette, A., Matise, T. C., McKenzie, C. A., McKnight, B., Moll, F. L., Morris, A. D., Morris, A. P., Murray, J. C., Nelis, M., Ohlsson, C., Oldehinkel, A. J., Ong, K. K., Madden, P. A. F., Pasterkamp, G., Peden, J. F., Peters, A., Postma, D. S., Pramstaller, P. P., Price, J. F., Qi, L., Raitakari, O. T., Rankinen, T., Rao, D. C., Rice, T. K., Ridker, P. M., Rioux, J. D., Ritchie, M. D., Rudan, I., Salomaa, V., Samani, N. J., Saramies, J., Sarzynski, M. A., Schunkert, H., Schwarz, P. E. H., Sever, P., Shuldiner, A. R., Sinisalo, J., Stolk, R. P., Strauch, K., Tönjes, A., Trégouët, D.-A., Tremblay, A., Tremoli, E., Virtamo, J., Vohl, M.-C., Völker, U., Waeber, G., Willemsen, G., Witteman, J. C., Zillikens, M. C., Adair, L. S., Amouyel, P., Asselbergs, F. W., Assimes, T. L., Bochud, M., Boehm, B. O., Boerwinkle, E., Bornstein, S. R., Bottinger, E. P., Bouchard, C., Cauchi, S., Chambers, J. C., Chanock, S. J., Cooper, R. S., de Bakker, P. I. W., Dedoussis, G.,

Ferrucci, L., Franks, P. W., Froguel, P., Groop, L. C., Haiman, C. A., Hamsten, A., Hui, J., Hunter, D. J., Hveem, K., Kaplan, R. C., Kivimaki, M., Kuh, D., Laakso, M., Liu, Y., Martin, N. G., März, W., Melbye, M., Metspalu, A., Moebus, S., Munroe, P. B., Njølstad, I., Oostra, B. A., Palmer, C. N. A., Pedersen, N. L., Perola, M., Pérusse, L., Peters, U., Power, C., Quertermous, T., Rauramaa, R., Rivadeneira, F., Saaristo, T. E., Saleheen, D., Sattar, N., Schadt, E. E., Schlessinger, D., Slagboom, P. E., Snieder, H., Spector, T. D., Thorsteinsdottir, U., Stumvoll, M., Tuomilehto, J., Uitterlinden, A. G., Uusitupa, M., van der Harst, P., Walker, M., Wallaschofski, H., Wareham, N. J., Watkins, H., Weir, D. R., Wichmann, H.-E., Wilson, J. F., Zanen, P., Borecki, I. B., Deloukas, P., Fox, C. S., Heid, I. M., O'Connell, J. R., Strachan, D. P., Stefansson, K., van Duijn, C. M., Abecasis, G. R., Franke, L., Frayling, T. M., McCarthy, M. I., Visscher, P. M., Scherag, A., Willer, C. J., Boehnke, M., Mohlke, K. L., Lindgren, C. M., Beckmann, J. S., Barroso, I., North, K. E., Ingelsson, E., Hirschhorn, J. N., Loos, R. J. F., and Speliotes, E. K. (2015). Genetic studies of body mass index yield new insights for obesity biology. *Nature*, 518(7538):197–206.

Lokaj-Berisha, V., Gacaferri-Lumezi, B., Minci-Bejtullahu, G., Latifi-Pupovci, H., Karahoda-Gjurgjeala, N., Berisha, N., and Morina, T. (2015). Gender Associated High Body Mass Index in Allergic Diseases. *Open access Macedonian journal of medical sciences*, 3(1):69–74.

Lu, Y. and Loos, R. J. F. (2013). Obesity genomics: assessing the transferability of susceptibility loci across diverse populations. *Genome Medicine*, 5(6):1–14.

MacKenzie, R. G. (2006). Obesity-associated mutations in the human melanocortin-4 receptor gene. *Peptides*, 27(2):395–403.

Maes, H. H., Neale, M. C., and Eaves, L. J. (1997). Genetic and environmental factors in relative body weight and human adiposity. *Behavior genetics*, 27(4):325–351.

Mali, P., Yang, L., Esvelt, K. M., Aach, J., Guell, M., DiCarlo, J. E., Norville, J. E., and Church, G. M. (2013). RNA-Guided Human Genome Engineering via Cas9. *Science*, 339(6121):823–826.

May, C. M., Doroszuk, A., and Zwaan, B. J. (2015). The effect of developmental nutrition on life span and fecundity depends on the adult reproductive environment in Drosophila melanogaster. *Ecology and evolution*, 5(6):1156–68.

Melcher, C., Bader, R., and Pankratz, M. J. (2007). Amino acids, taste circuits, and feeding behavior in Drosophila: towards understanding the psychology of feeding in flies and man. *The Journal of endocrinology*, 192(3):467–472.

Melcher, C., Bader, R., Walther, S., Simakov, O., and Pankratz, M. J. (2006). Neuromedin U and its putative Drosophila homolog hugin. *PLoS biology*, 4(3):e68.

Melcher, C. and Pankratz, M. J. (2005). Candidate gustatory interneurons modulating feeding behavior in the Drosophila brain. *PLoS biology*, 3(9):e305.

Meng, X., Wahlstrom, G., Immonen, T., Kolmer, M., Tirronen, M., Predel, R., Kalkkinen, N., Heino, T. I., Sariola, H., and Roos, C. (2002). The Drosophila hugin gene codes for myostimulatory and ecdysis-modifying neuropeptides. *Mechanisms of development*, 117(1-2):5–13.

Mertens, I., Meeusen, T., Huybrechts, R., De Loof, A., and Schoofs, L. (2002). Characterization of the short neuropeptide F receptor from Drosophila melanogaster. *Biochemical and biophysical research communications*, 297(5):1140–1148.

Michaud, J. L., Boucher, F., Melnyk, A., Gauthier, F., Goshu, E., Levy, E., Mitchell, G. A., Himms-Hagen, J., and Fan, C. M. (2001). Sim1 haploinsufficiency causes hyperphagia, obesity and reduction of the paraventricular nucleus of the hypothalamus. *Human molecular genetics*, 10(14):1465–1473.

Miyoshi, H., Souza, S. C., Zhang, H.-H., Strissel, K. J., Christoffolete, M. A., Kovsan, J., Rudich, A., Kraemer, F. B., Bianco, A. C., Obin, M. S., and Greenberg, A. S. (2006). Perilipin promotes hormone-sensitive lipase-mediated adipocyte lipolysis via phosphorylation-dependent and -independent mechanisms. *The Journal of biological chemistry*, 281(23):15837–15844.

Mizuta, E., Kokubo, Y., Yamanaka, I., Miyamoto, Y., Okayama, A., Yoshimasa, Y., Tomoike, H., Morisaki, H., and Morisaki, T. (2008). Leptin gene and leptin receptor gene polymorphisms are associated with sweet preference and obesity. *Hypertension research : official journal of the Japanese Society of Hypertension*, 31(6):1069–1077.

Mohr, S. E., Hu, Y., Kim, K., Housden, B. E., and Perrimon, N. (2014). Resources for Functional Genomics Studies in Drosophila melanogaster. *Genetics*, 197(1):1–18.

Moloney, A., Sattelle, D. B., Lomas, D. A., and Crowther, D. C. (2010). Alzheimer's disease: insights from Drosophila melanogaster models. *Trends in biochemical sciences*, 35(4):228–35.

Montague, C. T., Farooqi, I. S., Whitehead, J. P., Soos, M. A., Rau, H., Wareham, N. J., Sewter, C. P., Digby, J. E., Mohammed, S. N., Hurst, J. A., Cheetham, C. H., Earley, A. R., Barnett, A. H., Prins, J. B., and O'Rahilly, S. (1997). Congenital leptin deficiency is associated with severe early-onset obesity in humans. *Nature*, 387(6636):903–908.

Moore, M. (2010). The International Mouse Phenotyping Consortium Concise Business Plan. Technical report, IMPC.

Moran, T. H. (2000). Cholecystokinin and satiety: current perspectives. *Nutrition (Burbank, Los Angeles County, Calif.)*, 16(10):858–65.

Morris, A. P., Voight, B. F., Teslovich, T. M., Ferreira, T., Segrè, A. V., Steinthorsdottir, V., Strawbridge, R. J., Khan, H., Grallert, H., Mahajan, A., Prokopenko, I., Kang, H. M., Dina, C., Esko, T., Fraser, R. M., Kanoni, S., Kumar, A., Lagou, V., Langenberg, C., Luan, J., Lindgren, C. M., Müller-Nurasyid, M., Pechlivanis, S., Rayner, N. W., Scott, L. J., Wiltshire, S., Yengo, L., Kinnunen, L., Rossin, E. J., and Raychaudhuri, S. (2012a). Large-scale association analysis provides insights into the genetic architecture and pathophysiology of type 2 diabetes. *Nat Genet*, 44.

Morris, S. N. S., Coogan, C., Chamseddin, K., Fernandez-Kim, S. O., Kolli, S., Keller, J. N., and Bauer, J. H. (2012b). Development of diet-induced insulin resistance in adult Drosophila melanogaster. *Biochimica et Biophysica Acta (BBA) - Molecular Basis of Disease*, 1822(8):1230–1237.

Müller, M. J., Geisler, C., Blundell, J., Dulloo, A., Schutz, Y., Krawczak, M., Bosy-Westphal, A., Enderle, J., and Heymsfield, S. B. (2018). The case of GWAS of obesity: does body weight control play by the rules? *International Journal of Obesity*, 42(8):1395–1405.

Mullins, D., Kirby, D., Hwa, J., Guzzi, M., Rivier, J., and Parker, E. (2001). Identification of potent and selective neuropeptide Y Y(1) receptor agonists with orexigenic activity in vivo. *Molecular pharmacology*, 60(3):534–540.

Murillo-Maldonado, J. M., Sánchez-Chávez, G., Salgado, L. M., Salceda, R., and Riesgo-Escovar, J. R. (2011). Drosophila insulin pathway mutants affect visual physiology and brain function besides growth, lipid, and carbohydrate metabolism. *Diabetes*, 60(5):1632–6.

Musselman, L. P., Fink, J. L., Narzinski, K., Ramachandran, P. V., Hathiramani, S. S., Cagan, R. L., Baranski, T. J., Palanker Musselman, L., Fink, J. L., Narzinski, K., Ramachandran, P. V., Sukumar Hathiramani, S., Cagan, R. L., and Baranski, T. J. (2011). A high-sugar diet produces obesity and insulin resistance in wild-type Drosophila. *Dis Model Mech*, 4(6):842–849.

Musselman, L. P. and Kühnlein, R. P. (2018). Drosophila as a model to study obesity and metabolic disease. *The Journal of Experimental Biology*, 221(Suppl 1):jeb163881.

Must, A., Spadano, J., Coakley, E. H., Field, A. E., Colditz, G., and Dietz, W. H. (1999). The disease burden associated with overweight and obesity. *JAMA*, 282(16):1523–1529.

Mutch, D. M. and Clement, K. (2006). Unraveling the genetics of human obesity. *PLoS genetics*, 2(12):e188.

Na, J., Musselman, L. P., Pendse, J., Baranski, T. J., Bodmer, R., Ocorr, K., and Cagan, R. (2013). A Drosophila Model of High Sugar Diet-Induced Cardiomyopathy. *PLoS Genet*, 9(1):e1003175.

NamKoong, C., Kim, M. S., Jang, B.-T., Lee, Y. H., Cho, Y.-M., and Choi, H. J. (2017). Central administration of GLP-1 and GIP decreases feeding in mice. *Biochemical and Biophysical Research Communications*, 490(2):247–252.

Nässel, D. R. (2018). Substrates for Neuronal Cotransmission With Neuropeptides and Small Molecule Neurotransmitters in Drosophila. *Frontiers in Cellular Neuroscience*, 12:83.

Nassel, D. R. and Broeck, J. V. (2016). Insulin/IGF signaling in Drosophila and other insects: factors that regulate production, release and post-release action of the insulin-like peptides. *Cellular and Molecular Life Sciences*, 73(2):271–290.

NCD Risk Factor Collaboration (2016). Trends in adult body-mass index in 200 countries from 1975 to 2014: a pooled analysis of 1698 population-based measurement studies with 19.2 million participants. *The Lancet*, 387(10026):1377–1396.

Neckameyer, W. S. and Argue, K. J. (2013). Comparative approaches to the study of physiology: Drosophila as a physiological tool. *American Journal of Physiology - Regulatory, Integrative and Comparative Physiology*, 304(3):R177–88.

Nichols, C. D., Becnel, J., and Pandey, U. B. (2012). Methods to Assay Drosophila Behavior. *Journal of Visualized Experiments : JoVE*, 61.

Nicolae, D. L., Gamazon, E., Zhang, W., Duan, S., Dolan, M. E., and Cox, N. J. (2010). Trait-Associated SNPs Are More Likely to Be eQTLs: Annotation to Enhance Discovery from GWAS. *PLoS Genetics*, 6(4):e1000888.

Nies, V. J. M., Struik, D., Wolfs, M. G. M., Rensen, S. S., Szalowska, E., Unmehopa, U. A., Fluiter, K., van der Meer, T. P., Hajmousa, G., Buurman, W. A., Greve, J. W., Rezaee, F., Shiri-Sverdlov, R., Vonk, R. J., Swaab, D. F., Wolffenbuttel, B. H. R., Jonker, J. W., and van Vliet-Ostaptchouk, J. V. (2018). TUB gene expression in hypothalamus and adipose tissue and its association with obesity in humans. *International Journal of Obesity*, 42(3):376–383.

Nishi, M. and Nanjo, K. (2011). Insulin gene mutations and diabetes. *Journal of Diabetes Investigation*, 2(2):92–100.

Noben-Trauth, K., Naggert, J. K., North, M. A., and Nishina, P. M. (1996). A candidate gene for the mouse mutation tubby. *Nature*, 380(6574):534–538.

Okamura, T., Yanobu-Takanashi, R., Takeuchi, F., Isono, M., Akiyama, K., Shimizu, Y., Goto, M., Liang, Y.-Q., Yamamoto, K., Katsuya, T., Fujioka, A., Ohnaka, K., Takayanagi, R., Ogihara, T., Yamori, Y., and Kato, N. (2012). Deletion of CDKAL1 Affects High-Fat Diet–Induced Fat Accumulation and Glucose-Stimulated Insulin Secretion in Mice, Indicating Relevance to Diabetes. *PLoS ONE*, 7(11):e49055.

Oksanen, L., Kainulainen, K., Heiman, M., Mustajoki, P., Kauppinen-Makelin, R., and Kontula, K. (1997). Novel polymorphism of the human ob gene promoter in lean and morbidly obese subjects. *International journal of obesity and related metabolic disorders : journal of the International Association for the Study of Obesity*, 21(6):489–494.

Ollmann, M. M., Wilson, B. D., Yang, Y. K., Kerns, J. A., Chen, Y., Gantz, I., and Barsh, G. S. (1997). Antagonism of central melanocortin receptors in vitro and in vivo by agouti-related protein. *Science*, 278(5335):135–138.

Ormerod, K. G., LePine, O. K., Abbineni, P. S., Bridgeman, J. M., Coorssen, J. R., Mercier, A. J., and Tattersall, G. J. (2017). Drosophila development, physiology, behavior, and lifespan are influenced by altered dietary composition. *Fly*, 11(3):153–170.

Panariello, F., De Luca, V., and de Bartolomeis, A. (2011). Weight Gain, Schizophrenia and Antipsychotics: New Findings from Animal Model and Pharmacogenomic Studies. *Schizophrenia Research and Treatment*, 2011:1–16.

Pandey, U. B. and Nichols, C. D. (2011). Human Disease Models in Drosophila melanogaster and the Role of the Fly in Therapeutic Drug Discovery. *Pharmacological Reviews*, 63(2):411–436.

Pantavou, K. G., Braliou, G. G., Kontou, P. I., Dimou, N. L., and Bagos, P. G. (2016). A meta-analysis of FZD3 gene polymorphisms and their association with schizophrenia. *Psychiatric Genetics*, 26(6):272–280.

Papanastasiou, S. A., Nakas, C. T., Carey, J. R., and Papadopoulos, N. T. (2013). Condition-Dependent Effects of Mating on Longevity and Fecundity of Female Medflies: The Interplay between Nutrition and Age of Mating. *PLoS ONE*, 8(7):e70181.

Park, J. H., Hogrebe, M., Grüneberg, M., DuChesne, I., von der Heiden, A. L., Reunert, J., Schlingmann, K. P., Boycott, K. M., Beaulieu, C. L., Mhanni, A. A., Innes, A. M., Hörtnagel, K., Biskup, S., Gleixner, E. M., Kurlemann, G., Fiedler, B., Omran, H., Rutsch, F., Wada, Y., Tsiakas, K., Santer, R., Nebert, D. W., Rust, S., and Marquardt, T. (2015). SLC39A8 Deficiency: A Disorder of Manganese Transport and Glycosylation. *American journal of human genetics*, 97(6):894–903.

Park, S., Alfa, R. W., Topper, S. M., Kim, G. E. S., Kockel, L., and Kim, S. K. (2014). A Genetic Strategy to Measure Circulating Drosophila Insulin Reveals Genes Regulating Insulin Production and Secretion. *PLoS Genetics*, 10(8):e1004555.

Pasco, M. Y. and Leopold, P. (2012). High Sugar-induced insulin resistance in Drosophila relies on the lipocalin neural lazarillo. *PLoS One*, 7.

Paternoster, L., Standl, M., Waage, J., Baurecht, H., Hotze, M., Strachan, D. P., Curtin, J. A., Bønnelykke, K., Tian, C., Takahashi, A., Esparza-Gordillo, J., Alves, A. C., Thyssen, J. P., den Dekker, H. T., Ferreira, M. A., Altmaier, E., Sleiman, P. M., Xiao, F. L., Gonzalez, J. R., Marenholz, I., Kalb, B., Yanes, M. P., Xu, C.-J., Carstensen, L., Groen-Blokhuis, M. M., Venturini, C., Pennell, C. E., Barton, S. J., Levin, A. M., Curjuric, I., Bustamante, M., Kreiner-Møller, E., Lockett, G. A., Bacelis, J., Bunyavanich, S., Myers, R. A., Matanovic, A., Kumar, A., Tung, J. Y., Hirota, T., Kubo, M., McArdle, W. L., Henderson, A. J., Kemp, J. P., Zheng, J., Smith, G. D., Rüschendorf, F., Bauerfeind, A., Lee-Kirsch, M. A., Arnold, A., Homuth, G., Schmidt, C. O., Mangold, E., Cichon, S., Keil, T., Rodríguez, E., Peters, A., Franke, A., Lieb, W., Novak, N., Fölster-Holst, R., Horikoshi, M., Pekkanen, J., Sebert, S., Husemoen, L. L., Grarup, N., de Jongste, J. C., Rivadeneira, F., Hofman, A., Jaddoe, V. W., Pasmans, S. G., Elbert, N. J., Uitterlinden, A. G., Marks, G. B., Thompson, P. J., Matheson, M. C., Robertson, C. F., Australian Asthma Genetics Consortium (AAGC), A. A. G. C., Ried, J. S., Li, J., Zuo, X. B., Zheng, X. D., Yin, X. Y., Sun, L. D., McAleer, M. A., O'Regan, G. M., Fahy, C. M., Campbell, L. E., Macek, M., Kurek, M., Hu, D., Eng, C., Postma, D. S., Feenstra, B., Geller, F., Hottenga, J. J., Middeldorp, C. M., Hysi, P., Bataille, V., Spector, T., Tiesler, C. M., Thiering, E., Pahukasahasram, B., Yang, J. J., Imboden, M., Huntsman, S., Vilor-Tejedor, N., Relton, C. L., Myhre, R., Nystad, W., Custovic, A., Weiss, S. T., Meyers, D. A., Söderhäll, C., Melén, E., Ober, C., Raby, B. A., Simpson, A., Jacobsson, B., Holloway, J. W., Bisgaard, H., Sunyer, J., Hensch, N. M. P., Williams, L. K., Godfrey, K. M., Wang, C. A., Boomsma, D. I., Melbye, M., Koppelman, G. H., Jarvis, D., McLean, W. I., Irvine, A. D., Zhang, X. J., Hakonarson, H., Gieger, C., Burchard, E. G., Martin, N. G., Duijts, L., Linneberg, A., Jarvelin, M.-R., Noethen, M. M., Lau, S., Hübner, N., Lee, Y.-A., Tamari, M., Hinds, D. A., Glass, D., Brown, S. J., Heinrich, J., Evans, D. M., and Weidinger, S. (2015). Multi-ancestry genome-wide association study of 21,000 cases and 95,000 controls identifies new risk loci for atopic dermatitis. *Nature genetics*, 47(12):1449–1456.

Pattanayak, V., Lin, S., Guilinger, J. P., Ma, E., Doudna, J. A., and Liu, D. R. (2013). High-throughput profiling of off-target DNA cleavage reveals RNA-programmed Cas9 nuclease specificity. *Nature Biotechnology*, 31(9):839–843.

Pearce, L. R., Atanassova, N., Banton, M. C., Bottomley, B., van der Klaauw, A. A., Revelli, J.-P., Hendricks, A., Keogh, J. M., Henning, E., Doree, D., Jeter-Jones, S., Garg, S., Bochukova, E. G., Bounds, R., Ashford, S., Gayton, E., Hindmarsh, P. C., Shield, J. P. H., Crowne, E., Barford, D., Wareham, N. J., UK10K consortium, U., O'Rahilly, S., Murphy, M. P., Powell, D. R., Barroso, I., and Farooqi, I. S. (2013). KSR2 mutations are associated with obesity, insulin resistance, and impaired cellular fuel oxidation. *Cell*, 155(4):765–77.

Pendse, J., Ramachandran, P. V., Na, J., Narisu, N., Fink, J. L., Cagan, R. L., Collins, F. S., and Baranski, T. J. (2013). A Drosophila functional evaluation of candidates from human genome-wide association studies of type 2 diabetes and related metabolic traits identifies tissue-specific roles for dHHEX. *BMC Genomics*, 14(1):1–11.

Perkins, L. A., Holderbaum, L., Tao, R., Hu, Y., Sopko, R., McCall, K., Yang-Zhou, D., Flockhart, I., Binari, R., Shim, H.-S., Miller, A., Housden, A., Foos, M., Randkelv, S., Kelley, C., Namgyal, P., Villalta, C., Liu, L.-P., Jiang, X., Huan-Huan, Q., Wang, X., Fujiyama, A., Toyoda, A., Ayers, K., Blum, A., Czech, B., Neumuller, R., Yan, D., Cavallaro, A., Hibbard, K., Hall, D., Cooley, L., Hannon, G. J., Lehmann, R., Parks, A., Mohr, S. E., Ueda, R., Kondo, S., Ni, J.-Q., and Perrimon, N. (2015). The Transgenic RNAi Project at Harvard Medical School: Resources and Validation. *Genetics*, 201(3):843–52.

Peterfy, M., Phan, J., Xu, P., and Reue, K. (2001). Lipodystrophy in the fld mouse results from mutation of a new gene encoding a nuclear protein, lipin. *Nature genetics*, 27(1):121–124.

Phifer-Rixey, M. and Nachman, M. W. (2015). Insights into mammalian biology from the wild house mouse Mus musculus. *eLife*, 4.

Ponton, F., Chapuis, M.-P., Pernice, M., Sword, G. A., and Simpson, S. J. (2011). Evaluation of potential reference genes for reverse transcription-qPCR studies of physiological responses in Drosophila melanogaster. *Journal of insect physiology*, 57(6):840–850.

Pool, A.-H. and Scott, K. (2014). Feeding regulation in Drosophila. *Current opinion in neurobiology*, 29:57–63.

Port, F. and Bullock, S. L. (2016). Augmenting CRISPR applications in Drosophila with tRNA-flanked sgRNAs. *Nature methods*, 13(10):852–854.

Port, F., Chen, H.-M., Lee, T., and Bullock, S. L. (2014). Optimized CRISPR/Cas tools for efficient germline and somatic genome engineering in Drosophila. *Proceedings of the National Academy of Sciences of the United States of America*, 111(29):E2967–76.

Pospisilik, J. A., Schramek, D., Schnidar, H., Cronin, S. J. F., Nehme, N. T., Zhang, X., Knauf, C., Cani, P. D., Aumayr, K., Todoric, J., Bayer, M., Haschemi, A., Puviindran, V., Tar, K., Orthofer, M., Neely, G. G., Dietzl, G., Manoukian, A., Funovics, M., Prager, G., Wagner, O., Ferrandon, D., Aberger, F., Hui, C.-c., Esterbauer, H., and Penninger, J. M. (2010). Drosophila genome-wide obesity screen reveals hedgehog as a determinant of brown versus white adipose cell fate. *Cell*, 140(1):148–60.

Pralong, F. P., Gonzales, C., Voirol, M.-J., Palmiter, R. D., Brunner, H.-R., Gaillard, R. C., Seydoux, J., and Pedrazzini, T. (2002). The neuropeptide Y Y1 receptor regulates leptin-mediated control of energy homeostasis and reproductive functions. *FASEB journal : official publication of the Federation of American Societies for Experimental Biology*, 16(7):712–714.

Radford, J. C., Davies, S. A., and Dow, J. A. T. (2002). Systematic G-protein-coupled receptor analysis in Drosophila melanogaster identifies a leucokinin receptor with novel roles. *The Journal of biological chemistry*, 277(41):38810–38817.

Rajan, A. and Perrimon, N. (2012). Drosophila cytokine unpaired 2 regulates physiological homeostasis by remotely controlling insulin secretion. *Cell*, 151(1):123–137.

Rathjen, T., Yan, X., Kononenko, N. L., Ku, M.-C., Song, K., Ferrarese, L., Tarallo, V., Puchkov, D., Kochlamazashvili, G., Brachs, S., Varela, L., Szigeti-Buck, K., Yi, C.-X., Schriever, S. C., Tattikota, S. G., Carlo, A. S., Moroni, M., Siemens, J., Heuser, A., van der Weyden, L., Birkenfeld, A. L., Niendorf, T., Poulet, J. F. A., Horvath, T. L., Tschöp, M. H., Heinig, M., Trajkovski, M., Haucke, V., and Poy, M. N. (2017). Regulation of body weight and energy homeostasis by neuronal cell adhesion molecule 1. *Nature Neuroscience*, 20(8):1096–1103.

Reis, T., Van Gilst, M. R., and Hariharan, I. K. (2010). A buoyancy-based screen of Drosophila larvae for fat-storage mutants reveals a role for Sir2 in coupling fat storage to nutrient availability. *PLoS genetics*, 6(11):e1001206.

Reiter, L. T., Potocki, L., Chien, S., Gribskov, M., and Bier, E. (2001). A systematic analysis of human disease-associated gene sequences in Drosophila melanogaster. *Genome research*, 11(6):1114–1125.

Ren, D., Li, M., Duan, C., and Rui, L. (2005). Identification of SH2-B as a key regulator of leptin sensitivity, energy balance, and body weight in mice. *Cell Metabolism*, 2(2):95–104.

Reue, K. (2009). The Lipin Family: Mutations and Metabolism. *Current opinion in lipidology*, 20(3):165–170.

Riediger, T., Traebert, M., Schmid, H. A., Scheel, C., Lutz, T. A., and Scharrer, E. (2003). Site-specific effects of ghrelin on the neuronal activity in the hypothalamic arcuate nucleus. *Neuroscience letters*, 341(2):151–5.

Robinow, S. and White, K. (1991). Characterization and spatial distribution of the ELAV protein duringDrosophila melanogaster development. *Journal of Neurobiology*, 22(5):443–461.

Rooney, K. and Ozanne, S. E. (2011). Maternal over-nutrition and offspring obesity predisposition: targets for preventative interventions. *International Journal of Obesity*, 35(7):883–890.

Root, C. M., Ko, K. I., Jafari, A., and Wang, J. W. (2011). Presynaptic Facilitation by Neuropeptide Signaling Mediates Odor-Driven Food Search. *Cell*, 145(1):133–144.

Rosenkilde, C., Cazzamali, G., Williamson, M., Hauser, F., Sondergaard, L., DeLotto, R., and Grimmelikhuijzen, C. J. P. (2003). Molecular cloning, functional expression, and gene silencing of two Drosophila receptors for the Drosophila neuropeptide pyrokinin-2. *Biochemical and biophysical research communications*, 309(2):485–494.

Rosenthal, N. and Brown, S. (2007). The mouse ascending: perspectives for human-disease models. *Nature cell biology*, 9(9):993–999.

Rossetti, L., Stenbit, A. E., Chen, W., Hu, M., Barzilai, N., Katz, E. B., and Charron, M. J. (1997). Peripheral but not hepatic insulin resistance in mice with one disrupted allele of the glucose transporter type 4 (GLUT4) gene. *The Journal of clinical investigation*, 100(7):1831–9.

Rulifson, E. J., Kim, S. K., and Nusse, R. (2002). Ablation of insulin-producing neurons in flies: growth and diabetic phenotypes. *Science*, 296.

Sahu, A. (1998). Leptin decreases food intake induced by melanin-concentrating hormone (MCH), galanin (GAL) and neuropeptide Y (NPY) in the rat. *Endocrinology*, 139(11):4739–4742.

Sahu, A., Kalra, P. S., Dube, M. G., and Kalra, S. P. (1988). Neuropeptide K suppresses feeding in the rat. *Regulatory peptides*, 23(2):135–143.

Sainsbury, A., Cusin, I., Rohner-Jeanrenaud, F., and Jeanrenaud, B. (1997). Adrenalectomy prevents the obesity syndrome produced by chronic central neuropeptide Y infusion in normal rats. *Diabetes*, 46(2):209–214.

Saller, C. F. and Stricker, E. M. (1976). Hyperphagia and increased growth in rats after intraventricular injection of 5,7-dihydroxytryptamine. *Science (New York, N.Y.)*, 192(4237):385–387.

Scarborough, P., Bhatnagar, P., Wickramasinghe, K. K., Allender, S., Foster, C., and Rayner, M. (2011). The economic burden of ill health due to diet, physical inactivity, smoking, alcohol and obesity in the UK: an update to 2006-07 NHS costs. *Journal of public health (Oxford, England)*, 33(4):527–535.

Schoofs, A., Hückesfeld, S., Schlegel, P., Miroschnikow, A., Peters, M., Zeymer, M., Spieß, R., Chiang, A.-S., and Pankratz, M. J. (2014). Selection of motor programs for suppressing food intake and inducing locomotion in the Drosophila brain. *PLoS biology*, 12(6):e1001893.

Schwartz, M. W., Sipols, A. J., Marks, J. L., Sanacora, G., White, J. D., Scheurink, A., Kahn, S. E., Baskin, D. G., Woods, S. C., and Figlewicz, D. P. (1992). Inhibition of hypothalamic neuropeptide Y gene expression by insulin. *Endocrinology*, 130(6):3608–3616.

Schwasinger-Schmidt, T. E., Kachman, S. D., and Harshman, L. G. (2012). Evolution of starvation resistance in Drosophila melanogaster: measurement of direct and correlated responses to artificial selection. *Journal of evolutionary biology*, 25(2):378–87.

Schweiger, M., Lass, A., Zimmermann, R., Eichmann, T. O., and Zechner, R. (2009). Neutral lipid storage disease: genetic disorders caused by mutations in adipose triglyceride lipase/PNPLA2 or CGI-58/ABHD5. *American journal of physiology. Endocrinology and metabolism*, 297(2):E289–96.

Seinen, E., Burgerhof, J. G. M., Jansen, R. C., and Sibon, O. C. M. (2011). RNAi-induced off-target effects in Drosophila melanogaster: frequencies and solutions. *Briefings in functional genomics*, 10(4):206–214.

Shen, P. (2012). Preparing Drosophila larvae for feeding assays. *Cold Spring Harbor protocols*, 2012(5):pdb.prot069302.

Shiraiwa, T. and Carlson, J. R. (2007). Proboscis extension response (PER) assay in Drosophila. *Journal of visualized experiments : JoVE*, (3):193.

Shungin, D., Winkler, T. W., Croteau-Chonka, D. C., Ferreira, T., Locke, A. E., Magi, R., Strawbridge, R. J., Pers, T. H., Fischer, K., Justice, A. E., Workalemahu, T., Wu, J. M. W., Buchkovich, M. L., Heard-Costa, N. L., Roman, T. S., Drong, A. W., Song, C., Gustafsson, S., Day, F. R., Esko, T., Fall, T., Kutalik, Z., Luan, J., Randall, J. C., Scherag, A., Vedantam, S., Wood, A. R., Chen, J., Fehrmann, R., Karjalainen, J., Kahali, B., Liu, C.-T., Schmidt, E. M., Absher, D., Amin, N., Anderson, D., Beekman, M., Bragg-Gresham, J. L., Buyske, S., Demirkan, A., Ehret, G. B., Feitosa, M. F., Goel, A., Jackson, A. U., Johnson, T., Kleber, M. E., Kristiansson, K., Mangino, M., Mateo Leach, I., Medina-Gomez, C., Palmer, C. N. A. C. D., Pasko, D., Pechlivanis, S., Peters, M. J., Prokopenko, I., Stancakova, A., Ju Sung, Y., Tanaka, T., Teumer, A., Van Vliet-Ostaptchouk, J. V., Yengo, L., Zhang, W., Albrecht, E., Arnlov, J., Arscott, G. M., Bandinelli, S., Barrett, A., Bellis, C., Bennett, A. J., Berne, C., Bluher, M., Bohringer, S., Bonnet, F., Bottcher, Y., Bruinenberg, M., Carba, D. B., Caspersen, I. H., Clarke, R., Daw, E. W., Deelen, J., Deelman, E., Delgado, G., Doney, A. S. F., Eklund, N., Erdos, M. R., Estrada, K., Eury, E., Friedrich, N., Garcia, M. E., Giedraitis, V., Gigante, B., Go, A. S., Golay, A., Grallert, H., Grammer, T. B., Grassler, J., Grewal, J., Groves, C. J., Haller, T., Hallmans, G., Hartman, C. A., Hassinen, M., Hayward, C., Heikkila, K., Herzig, K.-H., Helmer, Q., Hillege, H. L., Holmen, O., Hunt, S. C., Isaacs, A., Ittermann, T., James, A. L., Johansson, I., Juliusdottir, T., Kalafati, I.-P., Kinnunen, L., Koenig, W., Kooner, I. K., Kratzer, W., Lamina, C., Leander, K., Lee, N. R., Lichtner, P., Lind, L., Lindstrom, J., Lobbens, S., Lorentzon, M., Mach, F., Magnusson, P. K. E., Mahajan, A., McArdle, W. L., Menni, C., Merger, S., Mihailov, E., Milani, L., Mills, R., Moayyeri, A., Monda, K. L., Mooijaart, S. P., Muhleisen, T. W., Mulas, A., Muller, G., Muller-Nurasyid, M., Nagaraja, R., Nalls, M. A., Narisu, N., Glorioso, N., Nolte, I. M., Olden, M., Rayner, N. W., Renstrom, F., Ried, J. S., Robertson, N. R., Rose, L. M., Sanna, S., Scharnagl, H., Scholtens, S., Sennblad, B., Seufferlein, T., Sitlani, C. M., Vernon Smith, A., Stirrups, K., Stringham, H. M., Sundstrom, J., Swertz, M. A., Swift, A. J., Syvanen, A.-C., Tayo, B. O., Thorand, B., Thorleifsson, G., Tomaschitz, A., Troffa, C., van Oort, F. V. A., Verweij, N., Vonk, J. M., Waite, L. L., Wennauer, R., Wilsgaard, T., Wojczynski, M. K., Wong, A., Zhang, Q., Hua Zhao, J., Brennan, E. P., Choi, M., Eriksson, P., Folkersen, L., Franco-Cereceda, A., Gharavi, A. G., Hedman, A. K., Hivert, M.-F., Huang, J., Kanoni, S., Karpe, F., Keildson, S., Kiryluk, K., Liang, L., Lifton, R. P., Ma, B., McKnight, A. J., McPherson, R., Metspalu, A., Min, J. L., Moffatt, M. F., Montgomery, G. W., Murabito, J. M., Nicholson, G., Nyholt, D. R., Olsson, C., Perry, J. R. B., Reinmaa, E., Salem, R. M., Sandholm, N., Schadt, E. E., Scott, R. A., Stolk, L., Vallejo, E. E., Westra, H.-J., Zondervan, K. T., Amouyel, P., Arveiler, D., Bakker, S. J. L., Beilby, J., Bergman, R. N., Blangero, J., Brown, M. J., Burnier, M., Campbell, H., Chakravarti, A., Chines, P. S., Claudi-Boehm, S., Collins, F. S., Crawford, D. C., Danesh, J., de Faire, U., de Geus, E. J. C., Dorr, M., Erbel, R., Eriksson, J. G., Farrall, M., Ferrannini, E., Ferrieres, J., Forouhi, N. G., Forrester, T., Franco, O. H., Gansevoort, R. T., Gieger, C., Gudnason, V., Haiman, C. A., Harris, T. B., Hattersley, A. T., Heliovaara, M., Hicks, A. A., Hingorani, A. D., Hoffmann, W., Hofman, A., Homuth, G., Humphries, S. E., Hypponen, E., Illig, T., Jarvelin, M.-R., Johansen, B., Jousilahti, P., Jula, A. M., Kaprio, J., Kee, F., Keinanen-Kiukaanniemi, S. M., Kooner, J. S., Kooperberg, C., Kovacs, P., Kraja, A. T., Kumari, M., Kuulasmaa, K., Kuusisto, J., Lakka, T. A., Langenberg, C., Le Marchand, L., Lehtimaki, T., Lyssenko, V., Mannisto, S., Marette, A., Matise, T. C., McKenzie, C. A., McKnight, B., Musk, A. W., Mohlenkamp, S., Morris, A. P. D., Nelis, M., Ohlsson, C., Oldehinkel, A. J.,

Ong, K. K., Palmer, L. J., Penninx, B. W., Peters, A., Pramstaller, P. P., Raitakari, O. T., Rankinen, T., Rao, D. C., Rice, T. K., Ridker, P. M., Ritchie, M. D., Rudan, I., Salomaa, V., Samani, N. J., Saramies, J., Sarzynski, M. A., Schwarz, P. E. H., Shuldiner, A. R., Staessen, J. A., Steinthorsdottir, V., Stolk, R. P., Strauch, K., Tonjes, A., Tremblay, A., Tremoli, E., Vohl, M.-C., Volker, U., Vollenweider, P., Wilson, J. F., Witteman, J. C., Adair, L. S., Bochud, M., Boehm, B. O., Bornstein, S. R., Bouchard, C., Cauchi, S., Caulfield, M. J., Chambers, J. C., Chasman, D. I., Cooper, R. S., Dedoussis, G., Ferrucci, L., Froguel, P., Grabe, H.-J., Hamsten, A., Hui, J., Hveem, K., Jockel, K.-H., Kivimaki, M., Kuh, D., Laakso, M., Liu, Y., Marz, W., Munroe, P. B., Njolstad, I., Oostra, B. A., Palmer, C. N. A. C. D., Pedersen, N. L., Perola, M., Perusse, L., Peters, U., Power, C., Quertermous, T., Rauramaa, R., Rivadeneira, F., Saaristo, T. E., Saleheen, D., Sinisalo, J., Slagboom, P. E., Snieder, H., Spector, T. D., Thorsteinsdottir, U., Stumvoll, M., Tuomilehto, J., Uitterlinden, A. G., Uusitupa, M., van der Harst, P., Veronesi, G., Walker, M., Wareham, N. J., Watkins, H., Wichmann, H.-E., Abecasis, G. R., Assimes, T. L., Berndt, S. I., Boehnke, M., Borecki, I. B., Deloukas, P., Franke, L., Frayling, T. M., Groop, L. C., Hunter, D. J., Kaplan, R. C., O'Connell, J. R., Qi, L., Schlessinger, D., Strachan, D. P., Stefansson, K., van Duijn, C. M., Willer, C. J., Visscher, P. M., Yang, J., Hirschhorn, J. N., Zillikens, M. C., McCarthy, M. I., Speliotes, E. K., North, K. E., Fox, C. S., Barroso, I., Franks, P. W., Ingelsson, E., Heid, I. M., Loos, R. J. F., Cupples, L. A., Morris, A. P. D., Lindgren, C. M., and Mohlke, K. L. (2015). New genetic loci link adipose and insulin biology to body fat distribution. *Nature*, 518(7538):187–196.

Siju, K. P., Bräcker, L. B., and Grunwald Kadow, I. C. (2014). Neural mechanisms of context-dependent processing of CO2 avoidance behavior in fruit flies. *Fly*, 8(2):68–74.

Simansky, K. J. (1996). Serotonergic control of the organization of feeding and satiety. *Behavioural brain research*, 73(1-2):37–42.

Simpson, K. A., Martin, N. M., and R. Bloom, S. (2009). Hypothalamic regulation of food intake and clinical therapeutic applications .

Sinkus, M. L., Graw, S., Freedman, R., Ross, R. G., Lester, H. A., and Leonard, S. (2015). The human CHRNA7 and CHRFAM7A genes: A review of the genetics, regulation, and function. *Neuropharmacology*, 96(Pt B):274–88.

Slack, C., Werz, C., Wieser, D., Alic, N., Foley, A., Stocker, H., Withers, D. J., Thornton, J. M., Hafen, E., and Partridge, L. (2010). Regulation of Lifespan, Metabolism, and Stress Responses by the Drosophila SH2B Protein, Lnk. *PLoS Genetics*, 6(3):e1000881.

Smith, S. J., Cases, S., Jensen, D. R., Chen, H. C., Sande, E., Tow, B., Sanan, D. A., Raber, J., Eckel, R. H., and Farese, R. V. (2000). Obesity resistance and multiple mechanisms of triglyceride synthesis in mice lacking Dgat. *Nature genetics*, 25(1):87–90.

Smith, W. W., Thomas, J., Liu, J., Li, T., and Moran, T. H. (2014). From fat fruit fly to human obesity. *Physiology & behavior*, 136:15–21.

Söderberg, J. A. E., Carlsson, M. A., and Nässel, D. R. (2012). Insulin-Producing Cells in the Drosophila Brain also Express Satiety-Inducing Cholecystokinin-Like Peptide, Drosulfakinin. *Frontiers in endocrinology*, 3:109.

Song, W., Ren, D., Li, W., Jiang, L., Cho, K. W., Huang, P., Fan, C., Song, Y., Liu, Y., and Rui, L. (2010). SH2B Regulation of Growth, Metabolism, and Longevity in Both Insects and Mammals. *Cell Metabolism*, 11(5):427–437.

Speakman, J. R. (2013). Functional analysis of seven genes linked to body mass index and adiposity by genome-wide association studies: a review. *Human heredity*, 75(2-4):57–79.

Speakman, J. R. (2015). The 'Fat Mass and Obesity Related' (FTO) gene: Mechanisms of Impact on Obesity and Energy Balance. *Current Obesity Reports*, 4(1):73–91.

Speakman, J. R., Loos, R. J. F., O'Rahilly, S., Hirschhorn, J. N., and Allison, D. B. (2018). GWAS for BMI: a treasure trove of fundamental insights into the genetic basis of obesity. *International Journal of Obesity*, 42(8):1524–1531.

Speliotes, E. K., Willer, C. J., Berndt, S. I., Monda, K. L., Thorleifsson, G., Jackson, A. U., Lango Allen, H., Lindgren, C. M., Luan, J., Magi, R., Randall, J. C., Vedantam, S., Winkler, T. W., Qi, L., Workalemahu, T., Heid, I. M., Steinthorsdottir, V., Stringham, H. M., Weedon, M. N., Wheeler, E., Wood, A. R., Ferreira, T., Weyant, R. J., Segre, A. V., Estrada, K., Liang, L., Nemesh, J., Park, J.-H., Gustafsson, S., Kilpelainen, T. O., Yang, J., Bouatia-Naji, N., Esko, T., Feitosa, M. F., Kutalik, Z., Mangino, M., Raychaudhuri, S., Scherag, A., Smith, A. V., Welch, R., Zhao, J. H., Aben, K. K., Absher, D. M., Amin, N., Dixon, A. L., Fisher, E., Glazer, N. L., Goddard, M. E., Heard-Costa, N. L., Hoesel, V., Hottenga, J.-J., Johansson, A., Johnson, T., Ketkar, S., Lamina, C., Li, S., Moffatt, M. F., Myers, R. H., Narisu, N., Perry, J. R. B., Peters, M. J., Preuss, M., Ripatti, S., Rivadeneira, F., Sandholt, C., Scott, L. J., Timpson, N. J., Tyrer, J. P., van Wingerden, S., Watanabe, R. M., White, C. C., Wiklund, F., Barlassina, C., Chasman, D. I., Cooper, M. N., Jansson, J.-O., Lawrence, R. W., Pellikka, N., Prokopenko, I., Shi, J., Thiering, E., Alavere, H., Alibrandi, M. T. S., Almgren, P., Arnold, A. M., Aspelund, T., Atwood, L. D., Balkau, B., Balmforth, A. J., Bennett, A. J., Ben-Shlomo, Y., Bergman, R. N., Bergmann, S., Biebermann, H., Blakemore, A. I. F., Boes, T., Bonnycastle, L. L., Bornstein, S. R., Brown, M. J., Buchanan, T. A., Busonero, F., Campbell, H., Cappuccio, F. P., Cavalcanti-Proenca, C., Chen, Y.-D. I., Chen, C.-M., Chines, P. S., Clarke, R., Coin, L., Connell, J., Day, I. N. M., den Heijer, M., Duan, J., Ebrahim, S., Elliott, P., Elosua, R., Eiriksdottir, G., Erdos, M. R., Eriksson, J. G., Facheris, M. F., Felix, S. B., Fischer-Posovszky, P., Folsom, A. R., Friedrich, N., Freimer, N. B., Fu, M., Gaget, S., Gejman, P. V., Geus, E. J. C., Gieger, C., Gjesing, A. P., Goel, A., Goyette, P., Grallert, H., Grassler, J., Greenawalt, D. M., Groves, C. J., Gudnason, V., Guiducci, C., Hartikainen, A.-L., Hassanali, N., Hall, A. S., Havulinna, A. S., Hayward, C., Heath, A. C., Hengstenberg, C., Hicks, A. A., Hinney, A., Hofman, A., Homuth, G., Hui, J., Igl, W., Iribarren, C., Isomaa, B., Jacobs, K. B., Jarick, I., Jewell, E., John, U., Jorgensen, T., Jousilahti, P., Jula, A., Kaakinen, M., Kajantie, E., Kaplan, L. M., Kathiresan, S., Kettunen, J., Kinnunen, L., Knowles, J. W., Kolcic, I., Konig, I. R., Koskinen, S., Kovacs, P., Kuusisto, J., Kraft, P., Kvaloy, K., Laitinen, J., Lantieri, O., Lanzani, C., Launer, L. J., Lecoeur, C., Lehtimaki, T., Lettre, G., Liu, J., Lokki, M.-L., Lorentzon, M., Luben, R. N., Ludwig, B., Manunta, P., Marek, D., Marre, M., Martin, N. G., McArdle, W. L., McCarthy, A., McKnight, B., Meitinger, T., Melander, O., Meyre, D., Midthjell, K., Montgomery, G. W., Morken, M. A., Morris, A. P., Mulic, R., Ngwa, J. S., Nelis, M., Neville, M. J., Nyholt, D. R., O'Donnell, C. J., O'Rahilly, S., Ong, K. K., Oostra, B., Pare, G., Parker, A. N., Perola, M., Pichler, I., Pietilainen, K. H., Platou, C. G. P., Polasek, O., Pouta, A., Rafelt, S., Raitakari, O., Rayner, N. W., Ridderstrale, M.,

Rief, W., Ruokonen, A., Robertson, N. R., Rzehak, P., Salomaa, V., Sanders, A. R., Sandhu, M. S., Sanna, S., Saramies, J., Savolainen, M. J., Scherag, S., Schipf, S., Schreiber, S., Schunkert, H., Silander, K., Sinisalo, J., Siscovick, D. S., Smit, J. H., Soranzo, N., Sovio, U., Stephens, J., Surakka, I., Swift, A. J., Tammesoo, M.-L., Tardif, J.-C., Teder-Laving, M., Teslovich, T. M., Thompson, J. R., Thomson, B., Tonjes, A., Tuomi, T., van Meurs, J. B. J., van Ommen, G.-J., Vatin, V., Viikari, J., Visvikis-Siest, S., Vitart, V., Vogel, C. I. G., Voight, B. F., Waite, L. L., Wallaschofski, H., Walters, G. B., Widen, E., Wiegand, S., Wild, S. H., Willemsen, G., Witte, D. R., Witteman, J. C., Xu, J., Zhang, Q., Zgaga, L., Ziegler, A., Zitting, P., Beilby, J. P., Farooqi, I. S., Hebebrand, J., Huikuri, H. V., James, A. L., Kahonen, M., Levinson, D. F., Macciardi, F., Nieminen, M. S., Ohlsson, C., Palmer, L. J., Ridker, P. M., Stumvoll, M., Beckmann, J. S., Boeing, H., Boerwinkle, E., Boomsma, D. I., Caulfield, M. J., Chanock, S. J., Collins, F. S., Cupples, L. A., Smith, G. D., Erdmann, J., Froguel, P., Gronberg, H., Gyllensten, U., Hall, P., Hansen, T., Harris, T. B., Hattersley, A. T., Hayes, R. B., Heinrich, J., Hu, F. B., Hveem, K., Illig, T., Jarvelin, M.-R., Kaprio, J., Karpe, F., Khaw, K.-T., Kiemeney, L. A., Krude, H., Laakso, M., Lawlor, D. A., Metspalu, A., Munroe, P. B., Ouwehand, W. H., Pedersen, O., Penninx, B. W., Peters, A., Pramstaller, P. P., Quertermous, T., Reinehr, T., Rissanen, A., Rudan, I., Samani, N. J., Schwarz, P. E. H., Shuldiner, A. R., Spector, T. D., Tuomilehto, J., Uda, M., Uitterlinden, A., Valle, T. T., Wabitsch, M., Waeber, G., Wareham, N. J., Watkins, H., Wilson, J. F., Wright, A. F., Zillikens, M. C., Chatterjee, N., McCarroll, S. A., Purcell, S., Schadt, E. E., Visscher, P. M., Assimes, T. L., Borecki, I. B., Deloukas, P., Fox, C. S., Groop, L. C., Haritunians, T., Hunter, D. J., Kaplan, R. C., Mohlke, K. L., O'Connell, J. R., Peltonen, L., Schlessinger, D., Strachan, D. P., van Duijn, C. M., Wichmann, H.-E., Frayling, T. M., Thorsteinsdottir, U., Abecasis, G. R., Barroso, I., Boehnke, M., Stefansson, K., North, K. E., McCarthy, M. I., Hirschhorn, J. N., Ingelsson, E., and Loos, R. J. F. (2010). Association analyses of 249,796 individuals reveal 18 new loci associated with body mass index. *Nature genetics*, 42(11):937–948.

St Johnston, D. (2002). The art and design of genetic screens: Drosophila melanogaster. *Nature reviews. Genetics*, 3(3):176–188.

Stenbit, A. E., Tsao, T. S., Li, J., Burcelin, R., Geenen, D. L., Factor, S. M., Houseknecht, K., Katz, E. B., and Charron, M. J. (1997). GLUT4 heterozygous knockout mice develop muscle insulin resistance and diabetes. *Nature medicine*, 3(10):1096–101.

Stephens, T. W., Basinski, M., Bristow, P. K., Bue-Valleskey, J. M., Burgett, S. G., Craft, L., Hale, J., Hoffmann, J., Hsiung, H. M., Kriauciunas, A., MacKellar, W., Rosteck, P. R., Schoner, B., Smith, D., Tinsley, F. C., Zhang, X.-Y., and Heiman, M. (1995). The role of neuropeptide Y in the antiobesity action of the obese gene product. *Nature*, 377(6549):530–2.

Stijnen, P., Ramos-Molina, B., O'Rahilly, S., and Creemers, J. W. M. (2016). PCSK1 Mutations and Human Endocrinopathies: From Obesity to Gastrointestinal Disorders. *Endocrine Reviews*, 37(4):347–371.

Strassburger, K. and Teleman, A. A. (2016). Protocols to Study Growth and Metabolism in Drosophila. In *Methods in molecular biology (Clifton, N.J.)*, volume 1478, pages 279–290.

Stunkard, A. J., Foch, T. T., and Hrubec, Z. (1986a). A twin study of human obesity. *JAMA*, 256(1):51–54.

Stunkard, A. J., Harris, J. R., Pedersen, N. L., and McClearn, G. E. (1990). The body-mass index of twins who have been reared apart. *The New England journal of medicine*, 322(21):1483–1487.

Stunkard, A. J., Sorensen, T. I., Hanis, C., Teasdale, T. W., Chakraborty, R., Schull, W. J., and Schulsinger, F. (1986b). An adoption study of human obesity. *The New England journal of medicine*, 314(4):193–198.

Sun, F. L., Cuaycong, M. H., Craig, C. A., Wallrath, L. L., Locke, J., and Elgin, S. C. (2000). The fourth chromosome of Drosophila melanogaster: interspersed euchromatic and heterochromatic domains. *Proceedings of the National Academy of Sciences of the United States of America*, 97(10):5340–5.

Sztalryd, C., Xu, G., Dorward, H., Tansey, J. T., Contreras, J. A., Kimmel, A. R., and Londos, C. (2003). Perilipin A is essential for the translocation of hormone-sensitive lipase during lipolytic activation. *The Journal of cell biology*, 161(6):1093–1103.

Tansey, J. T., Sztalryd, C., Gruia-Gray, J., Roush, D. L., Zee, J. V., Gavrilova, O., Reitman, M. L., Deng, C.-X., Li, C., Kimmel, A. R., and Londos, C. (2001). Perilipin ablation results in a lean mouse with aberrant adipocyte lipolysis, enhanced leptin production, and resistance to diet-induced obesity. *Proceedings of the National Academy of Sciences*, 98(11):6494 LP – 6499.

Tartaglia, L. A., Dembski, M., Weng, X., Deng, N., Culpepper, J., Devos, R., Richards, G. J., Campfield, L. A., Clark, F. T., Deeds, J., Muir, C., Sanker, S., Moriarty, A., Moore, K. J., Smutko, J. S., Mays, G. G., Wool, E. A., Monroe, C. A., and Tepper, R. I. (1995). Identification and expression cloning of a leptin receptor, OB-R. *Cell*, 83(7):1263–71.

Tatemoto, K., Carlquist, M., and Mutt, V. (1982). Neuropeptide Y–a novel brain peptide with structural similarities to peptide YY and pancreatic polypeptide. *Nature*, 296(5858):659–60.

Teixeira, L., Rabouille, C., Rorth, P., Ephrussi, A., and Vanzo, N. F. (2003). Drosophila Perilipin/ADRP homologue Lsd2 regulates lipid metabolism. *Mechanisms of development*, 120(9):1071–1081.

Tennessen, J. M., Barry, W. E., Cox, J., and Thummel, C. S. (2014). Methods for studying metabolism in Drosophila. *Methods (San Diego, Calif.)*, 68(1):105–115.

Terzian, T., Dumble, M., Arbab, F., Thaller, C., Donehower, L. A., Lozano, G., Justice, M. J., Roop, D. R., and Box, N. F. (2011). Rpl27a mutation in the sooty foot ataxia mouse phenocopies high p53 mouse models. *The Journal of Pathology*, 224(4):540–552.

Teubner, B. J. and Bartness, T. J. (2013). PYY(3-36) into the arcuate nucleus inhibits food deprivation-induced increases in food hoarding and intake. *Peptides*, 47:20–28.

Timmons, J. A., Atherton, P. J., Larsson, O., Sood, S., Blokhin, I. O., Brogan, R. J., Volmar, C.-H., Josse, A. R., Slentz, C., Wahlestedt, C., Phillips, S. M., Phillips, B. E., Gallagher, I. J., and Kraus, W. E. (2018). A coding and non-coding transcriptomic perspective on the genomics of human metabolic disease. *Nucleic Acids Research*.

Trindade de Paula, M., Poetini Silva, M. R., Machado Araujo, S., Cardoso Bortolotto, V., Barreto Meichtry, L., Zemolin, A. P. P., Wallau, G. L., Jesse, C. R., Franco, J. L., Posser, T., and Prigol, M. (2016). High-Fat Diet Induces Oxidative Stress and MPK2 and HSP83 Gene Expression in Drosophila melanogaster. *Oxidative medicine and cellular longevity*, 2016:4018157.

Tung, Y. L., Gulati, P., Liu, C.-H., Rimmington, D., Dennis, R., Ma, M., Saudek, V., O'Rahilly, S., Coll, A. P., and Yeo, G. S. (2015). FTO is necessary for the induction of leptin resistance by high-fat feeding. *Molecular Metabolism*, 4(4):287–298.

Ugrankar, R., Berglund, E., Akdemir, F., Tran, C., Kim, M. S., Noh, J., Schneider, R., Ebert, B., and Graff, J. M. (2015). Drosophila glucome screening identifies Ck1alpha as a regulator of mammalian glucose metabolism. *Nature communications*, 6:7102.

Ugrankar, R., Liu, Y., Provaznik, J., Schmitt, S., and Lehmann, M. (2011). Lipin is a central regulator of adipose tissue development and function in Drosophila melanogaster. *Molecular and cellular biology*, 31(8):1646–1656.

Van der Horst, D. J. (2003). Insect adipokinetic hormones: release and integration of flight energy metabolism. *Comparative biochemistry and physiology. Part B, Biochemistry & molecular biology*, 136(2):217–226.

Van Holde, K. E. K. E. and Zlatanova, J. (2018). *The evolution of molecular biology : the search for the secrets of life*. Academic Press, 1 edition.

van Rheenen, W., Shatunov, A., Dekker, A. M., McLaughlin, R. L., Diekstra, F. P., Pulit, S. L., van der Spek, R. A. A., Võsa, U., de Jong, S., Robinson, M. R., Yang, J., Fogh, I., van Doormaal, P. T., Tazelaar, G. H. P., Koppers, M., Blokhuis, A. M., Sproviero, W., Jones, A. R., Kenna, K. P., van Eijk, K. R., Harschnitz, O., Schellevis, R. D., Brands, W. J., Medic, J., Menelaou, A., Vajda, A., Ticozzi, N., Lin, K., Rogelj, B., Vrabec, K., Ravnik-Glavač, M., Koritnik, B., Zidar, J., Leonardis, L., Grošelj, L. D., Millecamps, S., Salachas, F., Meininger, V., de Carvalho, M., Pinto, S., Mora, J. S., Rojas-García, R., Polak, M., Chandran, S., Colville, S., Swingler, R., Morrison, K. E., Shaw, P. J., Hardy, J., Orrell, R. W., Pittman, A., Sidle, K., Fratta, P., Malaspina, A., Topp, S., Petri, S., Abdulla, S., Drepper, C., Sendtner, M., Meyer, T., Ophoff, R. A., Staats, K. A., Wiedau-Pazos, M., Lomen-Hoerth, C., Van Deerlin, V. M., Trojanowski, J. Q., Elman, L., McCluskey, L., Basak, A. N., Tunca, C., Hamzeiy, H., Parman, Y., Meitinger, T., Lichtner, P., Radivojkov-Blagojevic, M., Andres, C. R., Maurel, C., Bensimon, G., Landwehrmeyer, B., Brice, A., Payan, C. A. M., Saker-Delye, S., Dürr, A., Wood, N. W., Tittmann, L., Lieb, W., Franke, A., Rietschel, M., Cichon, S., Nöthen, M. M., Amouyel, P., Tzourio, C., Dartigues, J.-F., Uitterlinden, A. G., Rivadeneira, F., Estrada, K., Hofman, A., Curtis, C., Blauw, H. M., van der Kooi, A. J., de Visser, M., Goris, A., Weber, M., Shaw, C. E., Smith, B. N., Pansarasa, O., Cereda, C., Del Bo, R., Comi, G. P., D'Alfonso, S., Bertolin, C., Sorarù, G., Mazzini, L., Pensato, V., Gellera, C., Tiloca, C., Ratti, A., Calvo, A., Moglia, C., Brunetti, M., Arcuti, S., Capozzo, R., Zecca, C., Lunetta, C., Penco, S., Riva, N., Padovani, A., Filosto, M., Muller, B., Stuit, R. J., PARALS Registry, P., SLALOM Group, S., SLAP Registry, S., FALS Sequencing Consortium, F. S., SLAGEN Consortium, S., NNIPPS Study Group, N. S., Blair, I., Zhang, K., McCann, E. P., Fifita, J. A., Nicholson, G. A., Rowe, D. B., Pamphlett, R., Kiernan, M. C., Grosskreutz, J., Witte, O. W., Ringer, T., Prell, T., Stubendorff, B., Kurth, I., Hübner, C. A., Leigh, P. N., Casale, F., Chio, A.,

Beghi, E., Pupillo, E., Tortelli, R., Logroscino, G., Powell, J., Ludolph, A. C., Weishaupt, J. H., Robberecht, W., Van Damme, P., Franke, L., Pers, T. H., Brown, R. H., Glass, J. D., Landers, J. E., Hardiman, O., Andersen, P. M., Corcia, P., Vourc'h, P., Silani, V., Wray, N. R., Visscher, P. M., de Bakker, P. I. W., van Es, M. A., Pasterkamp, R. J., Lewis, C. M., Breen, G., Al-Chalabi, A., van den Berg, L. H., and Veldink, J. H. (2016). Genome-wide association analyses identify new risk variants and the genetic architecture of amyotrophic lateral sclerosis. *Nature genetics*, 48(9):1043–8.

Vandamme, T. F. (2014). Use of rodents as models of human diseases. *Journal of pharmacy & bioallied sciences*, 6(1):2–9.

Vang, L. L., Medvedev, A. V., and Adler, J. (2012). Simple Ways to Measure Behavioral Responses of Drosophila to Stimuli and Use of These Methods to Characterize a Novel Mutant. *PLoS ONE*, 7(5):e37495.

Varela, L. and Horvath, T. L. (2012). Leptin and insulin pathways in POMC and AgRP neurons that modulate energy balance and glucose homeostasis. *EMBO reports*, 13(12):1079–86.

VDRC (2016). www.vdrc.at.

Vissers, J. H., Manning, S. A., Kulkarni, A., and Harvey, K. F. (2016). A Drosophila RNAi library modulates Hippo pathway-dependent tissue growth. *Nature Communications*, 7:10368.

Wall, J. D. and Pritchard, J. K. (2003). Haplotype blocks and linkage disequilibrium in the human genome. *Nature Reviews Genetics*, 4(8):587–597.

Wang, J., Gong, J., Li, L., Chen, Y., Liu, L., Gu, H., Luo, X., Hou, F., Zhang, J., and Song, R. (2018). Neurexin gene family variants as risk factors for autism spectrum disorder. *Autism Research*, 11(1):37–43.

Wang, X., Rao, R. P., Kosakowska-Cholody, T., Masood, M. A., Southon, E., Zhang, H., Berthet, C., Nagashim, K., Veenstra, T. K., Tessarollo, L., Acharya, U., and Acharya, J. K. (2009). Mitochondrial degeneration and not apoptosis is the primary cause of embryonic lethality in ceramide transfer protein mutant mice. *The Journal of cell biology*, 184(1):143–58.

Wangler, M. F., Hu, Y., and Shulman, J. M. (2017). Drosophila and genome-wide association studies: a review and resource for the functional dissection of human complex traits. *Disease models & mechanisms*, 10(2):77–88.

Waterson, M. J., Chung, B. Y., Harvanek, Z. M., Ostojic, I., Alcedo, J., and Pletcher, S. D. (2014). Water sensor ppk28 modulates Drosophila lifespan and physiology through AKH signaling. *Proceedings of the National Academy of Sciences of the United States of America*, 111(22):8137–42.

Weissenburger, J., Rush, A. J., Giles, D. E., and Stunkard, A. J. (1986). Weight change in depression. *Psychiatry research*, 17(4):275–83.

Wheeler, E., Huang, N., Bochukova, E. G., Keogh, J. M., Lindsay, S., Garg, S., Henning, E., Blackburn, H., Loos, R. J. F., Wareham, N. J., O'Rahilly, S., Hurles, M. E., Barroso, I., and Farooqi, I. S. (2013). Genome-wide SNP and CNV analysis identifies common and low-frequency variants associated with severe early-onset obesity. *Nature genetics*, 45(5):513–7.

White, H., Pieper, C., Schmader, K., and Fillenbaum, G. (1996). Weight change in Alzheimer's disease. *Journal of the American Geriatrics Society*, 44(3):265–72.

WHO (2018). WHO fact sheet: Obesity and overweight. Technical report, WHO.

Williams, M. J., Eriksson, A., Shaik, M., Voisin, S., Yamskova, O., Paulsson, J., Thombare, K., Fredriksson, R., and Schiöth, H. B. (2015). The Obesity-Linked Gene Nudt3 Drosophila Homolog Aps Is Associated With Insulin Signaling. *Molecular endocrinology (Baltimore, Md.)*, 29(9):1303–19.

Williams, M. J., Goergen, P., Rajendran, J., Zheleznyakova, G., Hägglund, M. G., Perland, E., Bagchi, S., Kalogeropoulou, A., Khan, Z., Fredriksson, R., and Schiöth, H. B. (2014). Obesity-Linked Homologues TfAP-2 and Twz Establish Meal Frequency in Drosophila melanogaster. *PLoS Genetics*, 10(9):e1004499.

Winkelmann, J., Czamara, D., Schormair, B., Knauf, F., Schulte, E. C., Trenkwalder, C., Dauvilliers, Y., Polo, O., Högl, B., Berger, K., Fuhs, A., Gross, N., Stiasny-Kolster, K., Oertel, W., Bachmann, C. G., Paulus, W., Xiong, L., Montplaisir, J., Rouleau, G. A., Fietze, I., Vávrová, J., Kemlink, D., Sonka, K., Nevsimalova, S., Lin, S.-C., Wszolek, Z., Vilariño-Güell, C., Farrer, M. J., Gschliesser, V., Frauscher, B., Falkenstetter, T., Poewe, W., Allen, R. P., Earley, C. J., Ondo, W. G., Le, W.-D., Spieler, D., Kaffe, M., Zimprich, A., Kettunen, J., Perola, M., Silander, K., Cournu-Rebeix, I., Francavilla, M., Fontenille, C., Fontaine, B., Vodicka, P., Prokisch, H., Lichtner, P., Peppard, P., Faraco, J., Mignot, E., Gieger, C., Illig, T., Wichmann, H.-E., Müller-Myhsok, B., and Meitinger, T. (2011). Genome-Wide Association Study Identifies Novel Restless Legs Syndrome Susceptibility Loci on 2p14 and 16q12.1. *PLoS Genetics*, 7(7):e1002171.

Wittkopp, P. J., Haerum, B. K., and Clark, A. G. (2006). Parent-of-origin effects on mRNA expression in Drosophila melanogaster not caused by genomic imprinting. *Genetics*, 173(3):1817–21.

Wolins, N. E., Brasaemle, D. L., and Bickel, P. E. (2006). A proposed model of fat packaging by exchangeable lipid droplet proteins. *FEBS letters*, 580(23):5484–5491.

Wong, R., Piper, M. D. W., Wertheim, B., and Partridge, L. (2009). Quantification of food intake in Drosophila. *PloS one*, 4(6):e6063.

Wright, V. M., Vogt, K. L., Smythe, E., and Zeidler, M. P. (2011). Differential activities of the Drosophila JAK/STAT pathway ligands Upd, Upd2 and Upd3. *Cellular signalling*, 23(5):920–927.

Wu, Q. and Palmiter, R. D. (2011). GABAergic signaling by AgRP neurons prevents anorexia via a melanocortin-independent mechanism. *European Journal of Pharmacology*, 660(1):21–27.

Wu, Q., Wen, T., Lee, G., Park, J. H., Cai, H. N., and Shen, P. (2003). Developmental control of foraging and social behavior by the Drosophila neuropeptide Y-like system. *Neuron*, 39(1):147–161.

Wu, Q., Zhang, Y., Xu, J., and Shen, P. (2005a). Regulation of hunger-driven behaviors by neural ribosomal S6 kinase in Drosophila. *Proceedings of the National Academy of Sciences of the United States of America*, 102(37):13289–13294.

Wu, Q., Zhao, Z., and Shen, P. (2005b). Regulation of aversion to noxious food by Drosophila neuropeptide Y- and insulin-like systems. *Nature neuroscience*, 8(10):1350–1355.

Wu, Y.-W., Prakash, K. M., Rong, T.-Y., Li, H.-H., Xiao, Q., Tan, L. C., Au, W.-L., Ding, J.-q., Chen, S.-d., and Tan, E.-K. (2011). Lingo2 variants associated with essential tremor and Parkinson's disease. *Human Genetics*, 129(6):611–615.

Xia, Q. and Grant, S. F. A. (2013). The genetics of human obesity. *Annals of the New York Academy of Sciences*, 1281(1):178–190.

Xu, B., Goulding, E. H., Zang, K., Cepoi, D., Cone, R. D., Jones, K. R., Tecott, L. H., and Reichardt, L. F. (2003). Brain-derived neurotrophic factor regulates energy balance downstream of melanocortin-4 receptor. *Nature Neuroscience*, 6(7):736–742.

Xu, K., Zheng, X., and Sehgal, A. (2008). Regulation of feeding and metabolism by neuronal and peripheral clocks in Drosophila. *Cell metabolism*, 8(4):289–300.

Xu, W., Tan, L., and Yu, J.-T. (2015). The Role of PICALM in Alzheimer's Disease. *Molecular Neurobiology*, 52(1):399–413.

Yamada, K., Iwayama, Y., Hattori, E., Iwamoto, K., Toyota, T., Ohnishi, T., Ohba, H., Maekawa, M., Kato, T., and Yoshikawa, T. (2011). Genome-Wide Association Study of Schizophrenia in Japanese Population. *PLoS ONE*, 6(6):e20468.

Yan, X., Wang, Z., Schmidt, V., Gauert, A., Willnow, T. E., Heinig, M., and Poy, M. N. (2018). Cadm2 regulates body weight and energy homeostasis in mice. *Molecular Metabolism*, 8:180–188.

Yang, Z., Yu, Y., Zhang, V., Tian, Y., Qi, W., and Wang, L. (2015). Octopamine mediates starvation-induced hyperactivity in adult Drosophila. *Proceedings of the National Academy of Sciences of the United States of America*, 112(16):5219–5224.

Yazdi, F. T., Clee, S. M., and Meyre, D. (2015). Obesity genetics in mouse and human: back and forth, and back again. *PeerJ*, 3:e856.

Yen, C.-L. E., Stone, S. J., Koliwad, S., Harris, C., and Farese, R. V. (2008). DGAT enzymes and triacylglycerol biosynthesis. *Journal of Lipid Research*, 49(11):2283–2301.

Yeo, G. S. H., Connie Hung, C.-C., Rochford, J., Keogh, J., Gray, J., Sivaramakrishnan, S., O'Rahilly, S., and Farooqi, I. S. (2004). A de novo mutation affecting human TrkB associated with severe obesity and developmental delay. *Nature Neuroscience*, 7(11):1187–1189.

Young, F. B., Butland, S. L., Sanders, S. S., Sutton, L. M., and Hayden, M. R. (2012). Putting proteins in their place: Palmitoylation in Huntington disease and other neuropsychiatric diseases. *Progress in Neurobiology*, 97(2):220–238.

Zaltsman, Y., Shachnai, L., Yivgi-Ohana, N., Schwarz, M., Maryanovich, M., Houtkooper, R. H., Vaz, F. M., De Leonardis, F., Fiermonte, G., Palmieri, F., Gillissen, B., Daniel, P. T., Jimenez, E., Walsh, S., Koehler, C. M., Roy, S. S., Walter, L., Hajnóczky, G., and Gross, A. (2010). MTCH2/MIMP is a major facilitator of tBID recruitment to mitochondria. *Nature Cell Biology*, 12(6):553–562.

Zandawala, M., Yurgel, M. E., Liao, S., Texada, M. J., Rewitz, K. F., Keene, A. C., and Nassel, D. R. (2018). Orchestration of Drosophila post-feeding physiology and behavior by the neuropeptide leucokinin. *bioRxiv*.

Zerbino, D. R., Achuthan, P., Akanni, W., Amode, M. R., Barrell, D., Bhai, J., Billis, K., Cummins, C., Gall, A., Girón, C. G., Gil, L., Gordon, L., Haggerty, L., Haskell, E., Hourlier, T., Izuogu, O. G., Janacek, S. H., Juettemann, T., To, J. K., Laird, M. R., Lavidas, I., Liu, Z., Loveland, J. E., Maurel, T., McLaren, W., Moore, B., Mudge, J., Murphy, D. N., Newman, V., Nuhn, M., Ogeh, D., Ong, C. K., Parker, A., Patricio, M., Riat, H. S., Schuilenburg, H., Sheppard, D., Sparrow, H., Taylor, K., Thormann, A., Vullo, A., Walts, B., Zadissa, A., Frankish, A., Hunt, S. E., Kostadima, M., Langridge, N., Martin, F. J., Muffato, M., Perry, E., Ruffier, M., Staines, D. M., Trevanion, S. J., Aken, B. L., Cunningham, F., Yates, A., and Flicek, P. (2018). Ensembl 2018. *Nucleic Acids Research*, 46(D1):D754–D761.

Zhang, A. and Silverberg, J. I. (2015). Association of atopic dermatitis with being overweight and obese: a systematic review and metaanalysis. *Journal of the American Academy of Dermatology*, 72(4):606–16.e4.

Zhang, H., Liu, J., Li, C. R., Momen, B., Kohanski, R. A., and Pick, L. (2009). Deletion of Drosophila insulin-like peptides causes growth defects and metabolic abnormalities. *Proceedings of the National Academy of Sciences of the United States of America*, 106(46):19617–19622.

Zhang, J.-P., Li, X.-L., Neises, A., Chen, W., Hu, L.-P., Ji, G.-Z., Yu, J.-Y., Xu, J., Yuan, W.-P., Cheng, T., and Zhang, X.-B. (2016). Different Effects of sgRNA Length on CRISPR-mediated Gene Knockout Efficiency. *Scientific Reports*, 6(1):28566.

Zhang, L., Jia, R., Palange, N. J., Satheka, A. C., Togo, J., An, Y., Humphrey, M., Ban, L., Ji, Y., Jin, H., Feng, X., and Zheng, Y. (2015a). Large Genomic Fragment Deletions and Insertions in Mouse Using CRISPR/Cas9. *PLoS ONE*, 10(3).

Zhang, X.-H., Tee, L. Y., Wang, X.-G., Huang, Q.-S., and Yang, S.-H. (2015b). Off-target Effects in CRISPR/Cas9-mediated Genome Engineering. *Molecular Therapy - Nucleic Acids*, 4.

Zhang, Y., Proenca, R., Maffei, M., Barone, M., Leopold, L., and Friedman, J. M. (1994). Positional cloning of the mouse obese gene and its human homologue. *Nature*, 372(6505):425–432.

Zheng, H., Forgetta, V., Hsu, Y., Estrada, K., Rosello-Diez, A., Leo, P. J., Dahia, C. L., Park-Min, K. H., Tobias, J. H., Kooperberg, C., Kleinman, A., Styrkarsdottir, U., Liu, C., Uggla, C., Evans, D. S., Nielson, C. M., Walter, K., Pettersson-Kymmer, U., McCarthy, S., Eriksson, J., Kwan, T., Jhamai, M., Trajanoska, K., Memari, Y., Min, J., Huang, J., Danecek, P., Wilmot, B., Li, R., Chou, W., Mokry, L. E., Moayyeri, A., Claussnitzer, M., Cheng, C., Cheung, W., Medina-Gómez, C., Ge, B., Chen, S., Choi, K., Oei, L., Fraser, J., Kraaij, R., Hibbs, M. A., Gregson, C. L., Paquette, D., Hofman, A., Wibom, C., Tranah, G. J., Marshall, M., Gardiner, B. B., Cremin, K., Auer, P., Hsu, L., Ring, S., Tung, J. Y., Thorleifsson, G., Enneman, A. W., van Schoor, N. M., de Groot, L. C. P. G. M., van der Velde, N., Melin, B., Kemp, J. P., Christiansen, C., Sayers, A., Zhou, Y., Calderari, S., van Rooij, J., Carlson, C., Peters, U., Berlivet, S., Dostie, J., Uitterlinden, A. G., Williams, S. R., Farber, C., Grinberg, D., LaCroix, A. Z., Haessler, J., Chasman, D. I., Giulianini, F., Rose, L. M., Ridker, P. M., Eisman, J. A., Nguyen, T. V., Center, J. R., Nogues, X., Garcia-Giralt, N., Launer, L. L., Gudnason, V., Mellström, D., Vandenput, L., Amin, N., van Duijn, C. M., Karlsson, M. K., Ljunggren, Ö., Svensson, O., Hallmans, G., Rousseau, F., Giroux, S., Bussière, J., Arp, P. P., Koromani, F., Prince, R. L., Lewis, J. R., Langdahl, B. L., Pernille Hermann, A., Jensen, J. B., Kaptoge, S., Khaw, K., Reeve, J., Formosa, M. M., Xuereb-Anastasi, A., Åkesson, K., McGuigan, F. E., Garg, G., Olmos, J. M., Zarrabeitia, M. T., Riancho, J. A., Ralston, S. H., Alonso, N., Jiang, X., Goltzman, D., Pastinen, T., Grundberg, E., Gauguier, D., Orwoll, E. S., Karasik, D., Davey-Smith, G., Smith, A. V., Siggeirsdottir, K., Harris, T. B., Carola Zillikens, M., van Meurs, J. B. J., Thorsteinsdottir, U., Maurano, M. T., Timpson, N. J., Soranzo, N., Durbin, R., Wilson, S. G., Ntzani, E. E., Brown, M. A., Stefansson, K., Hinds, D. A., Spector, T., Adrienne Cupples, L., Ohlsson, C., Greenwood, C. M. T., Jackson, R. D., Rowe, D. W., Loomis, C. A., Evans, D. M., Ackert-Bicknell, C. L., Joyner, A. L., Duncan, E. L., Kiel, D. P., Rivadeneira, F., Richards, J. B., Rivadeneira, F., and Richards, J. B. (2015). Whole-genome sequencing identifies EN1 as a determinant of bone density and fracture. *Nature*, 526(7571):112–117.

Zhu, X., Zhou, A., Dey, A., Norrbom, C., Carroll, R., Zhang, C., Laurent, V., Lindberg, I., Ugleholdt, R., Holst, J. J., and Steiner, D. F. (2002). Disruption of PC1/3 expression in mice causes dwarfism and multiple neuroendocrine peptide processing defects. *Proceedings of the National Academy of Sciences*, 99(16):10293–10298.

Zinke, I., Kirchner, C., Chao, L. C., Tetzlaff, M. T., and Pankratz, M. J. (1999). Suppression of food intake and growth by amino acids in Drosophila: the role of pumpless, a fat body expressed gene with homology to vertebrate glycine cleavage system. *Development (Cambridge, England)*, 126(23):5275–5284.

Zisman, A., Peroni, O. D., Abel, E. D., Michael, M. D., Mauvais-Jarvis, F., Lowell, B. B., Wojtaszewski, J. F., Hirshman, M. F., Virkamaki, A., Goodyear, L. J., Kahn, C. R., and Kahn, B. B. (2000). Targeted disruption of the glucose transporter 4 selectively in muscle causes insulin resistance and glucose intolerance. *Nature medicine*, 6(8):924–8.

Zygmunt, D. A., Singhal, N., Kim, M.-L., Cramer, M. L., Crowe, K. E., Xu, R., Jia, Y., Adair, J., Martinez-Pena y Valenzuela, I., Akaaboune, M., White, P., Janssen, P. M., and Martin, P. T. (2017). Deletion of Pofut1 in Mouse Skeletal Myofibers Induces Muscle Aging-Related Phenotypes in cis and in trans. *Molecular and Cellular Biology*, 37(10).

Appendix A

Supplementary phenotype data

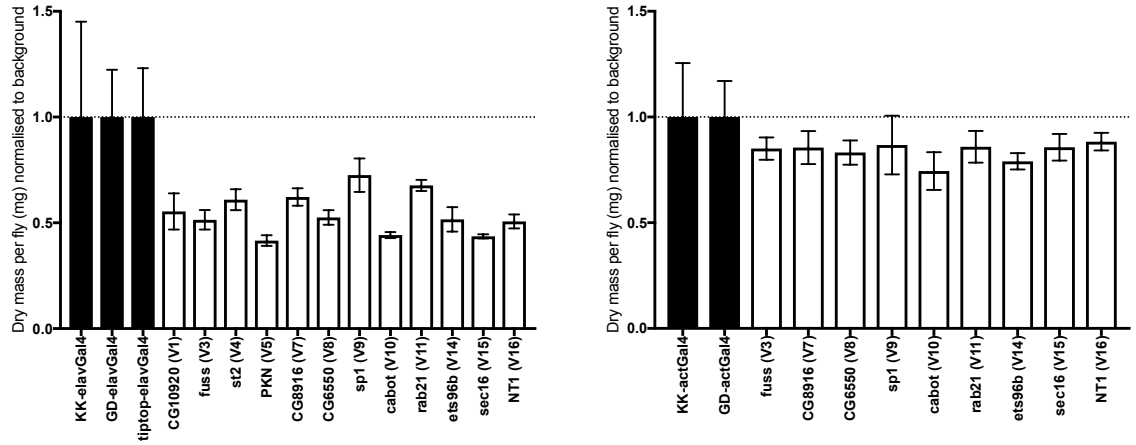


Fig. A.1 Dry mass of flies with neuronal (left) or whole-body (right) RNAi of BMI GWAS genes.

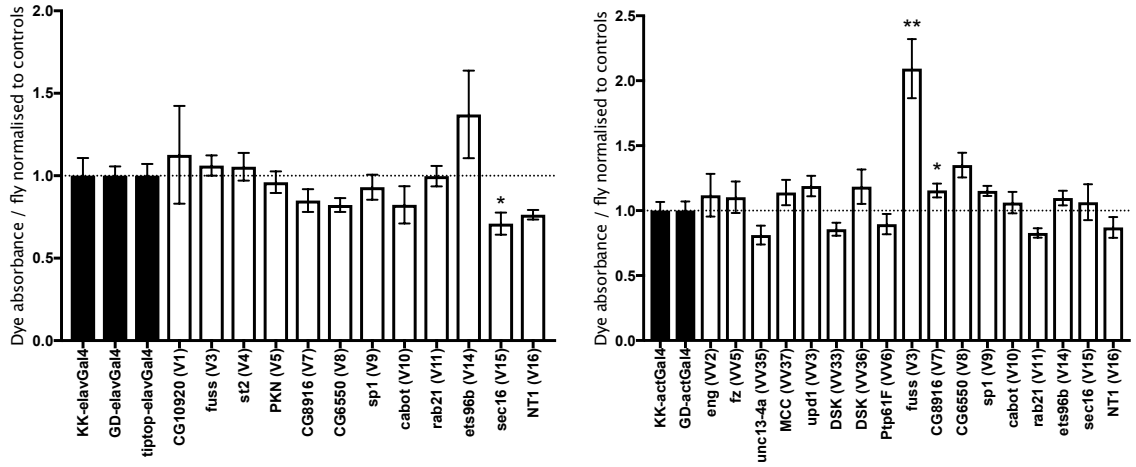


Fig. A.2 Dye absorbance of flies with neuronal (left) or whole-body (right) RNAi of BMI GWAS genes.

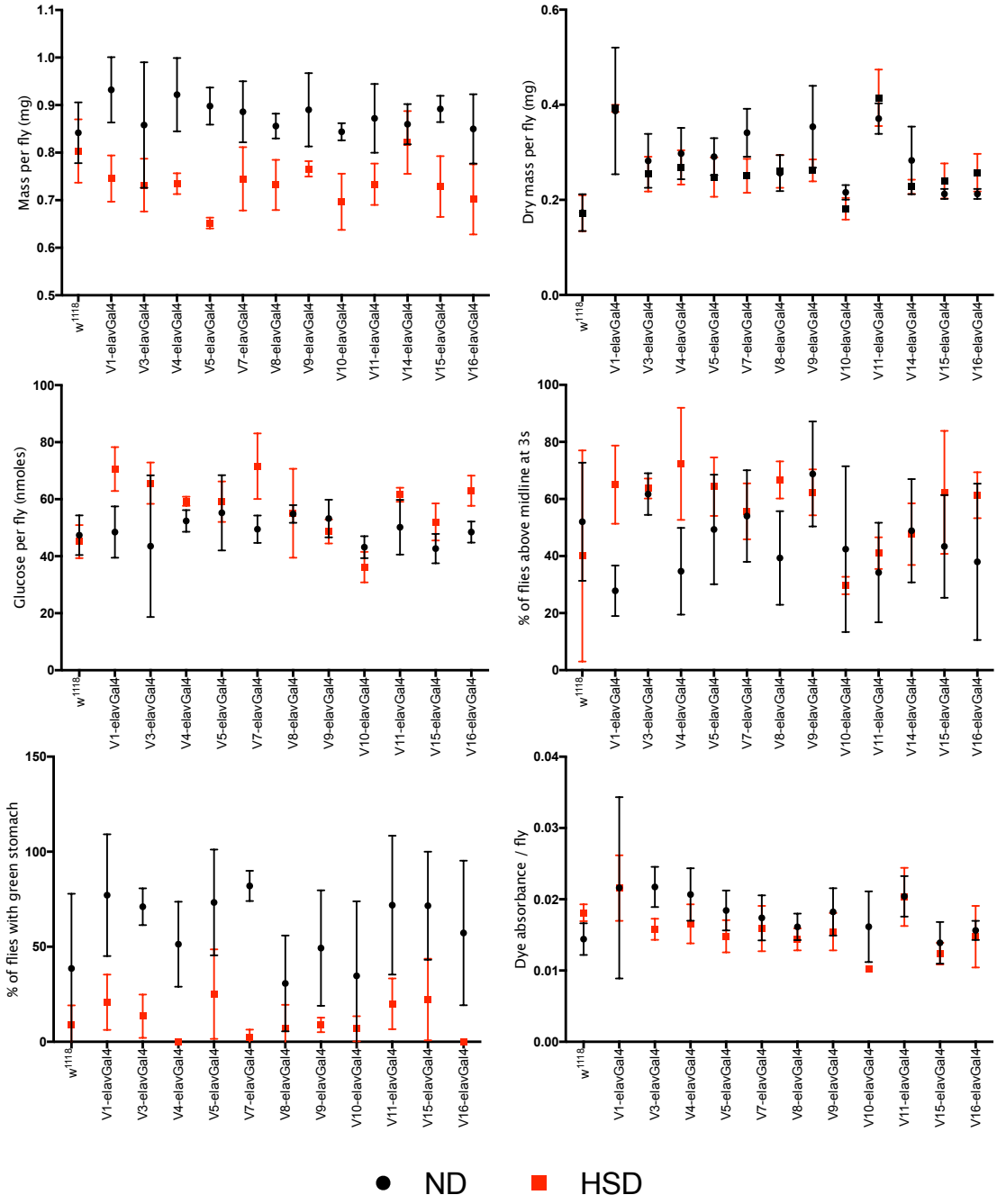


Fig. A.3 Phenotype of flies fed HSD compared to those fed ND.

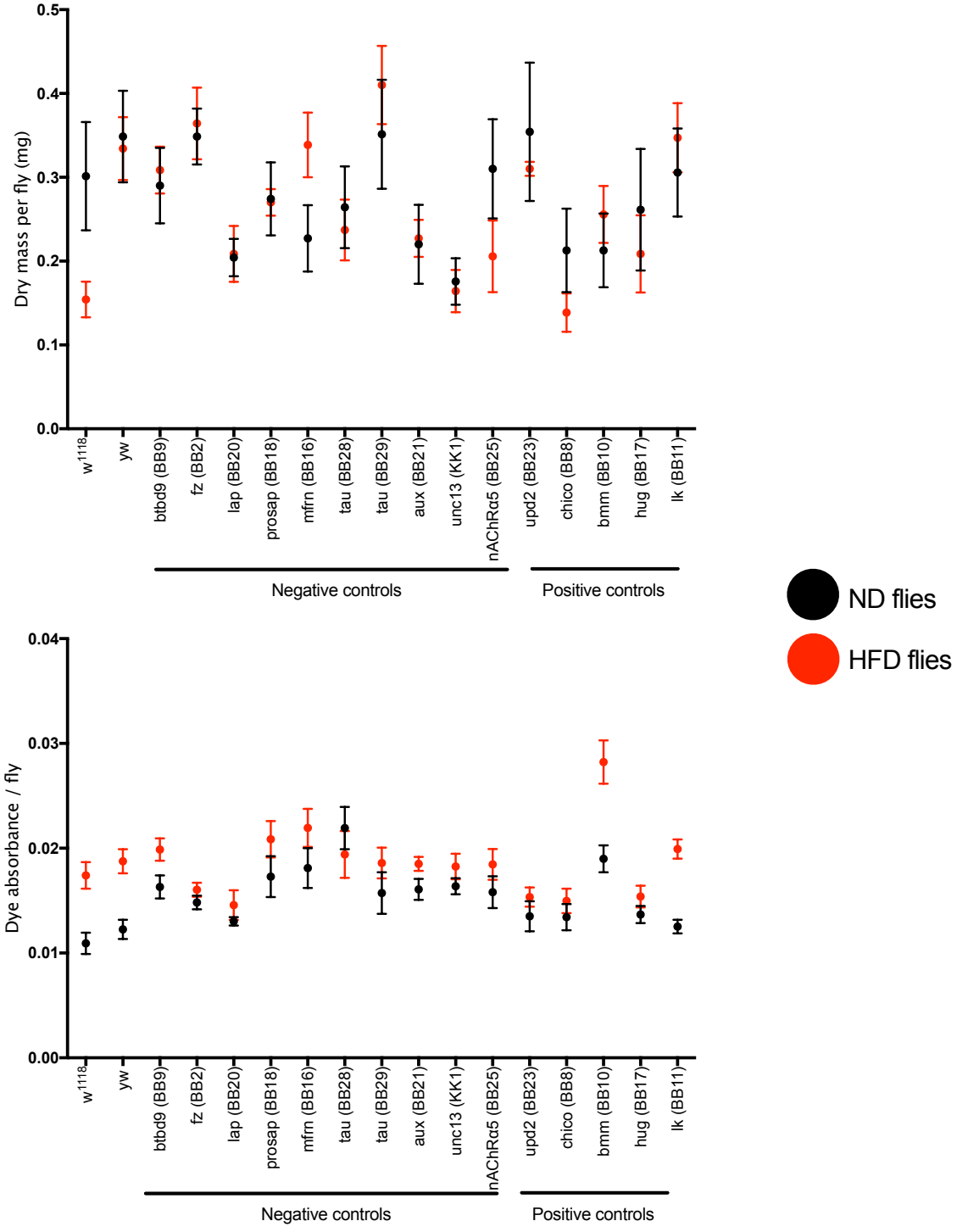


Fig. A.4 Phenotype of flies fed HFD compared to those fed ND.

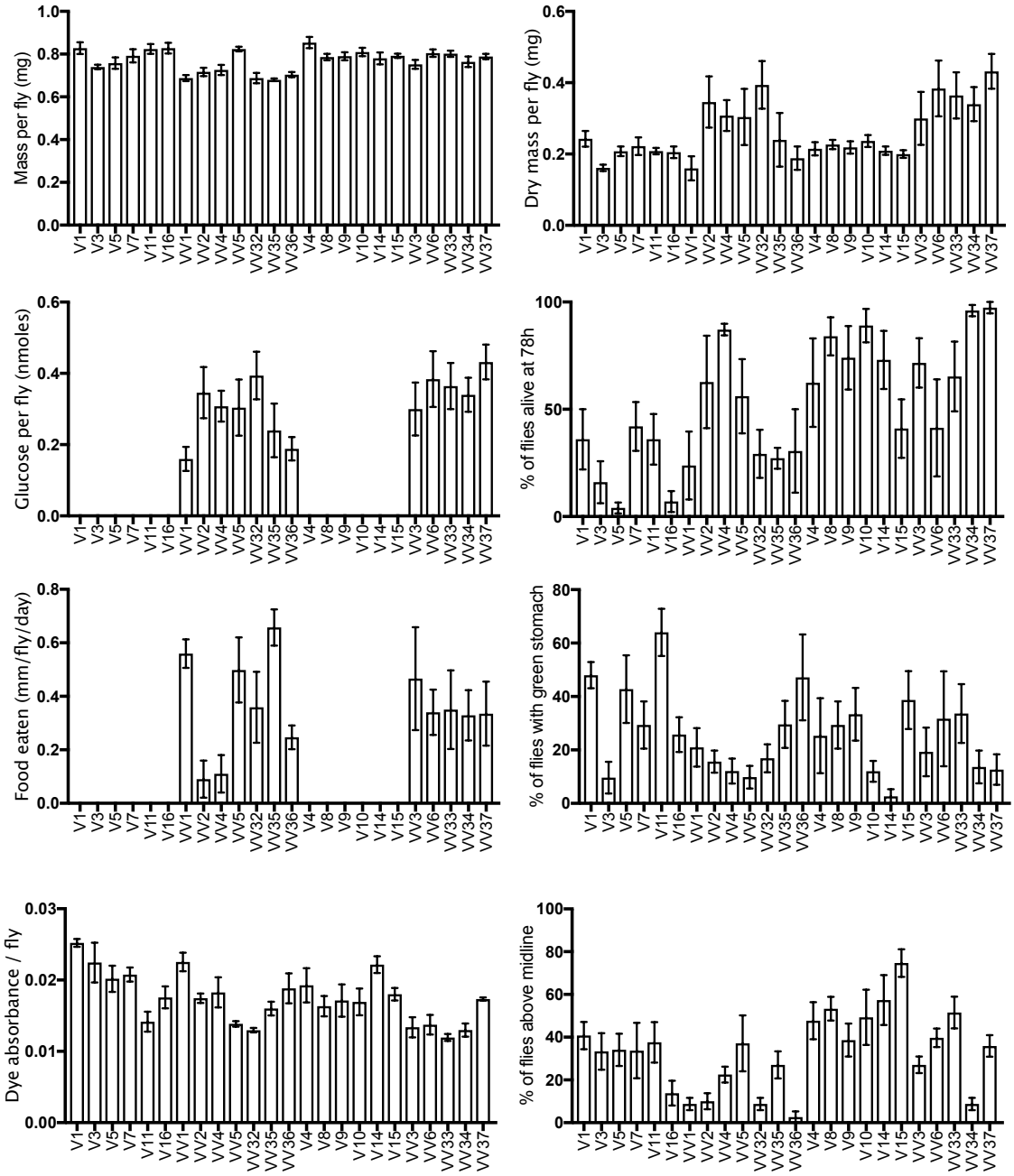


Fig. A.5 Phenotypes of parental siRNA lines

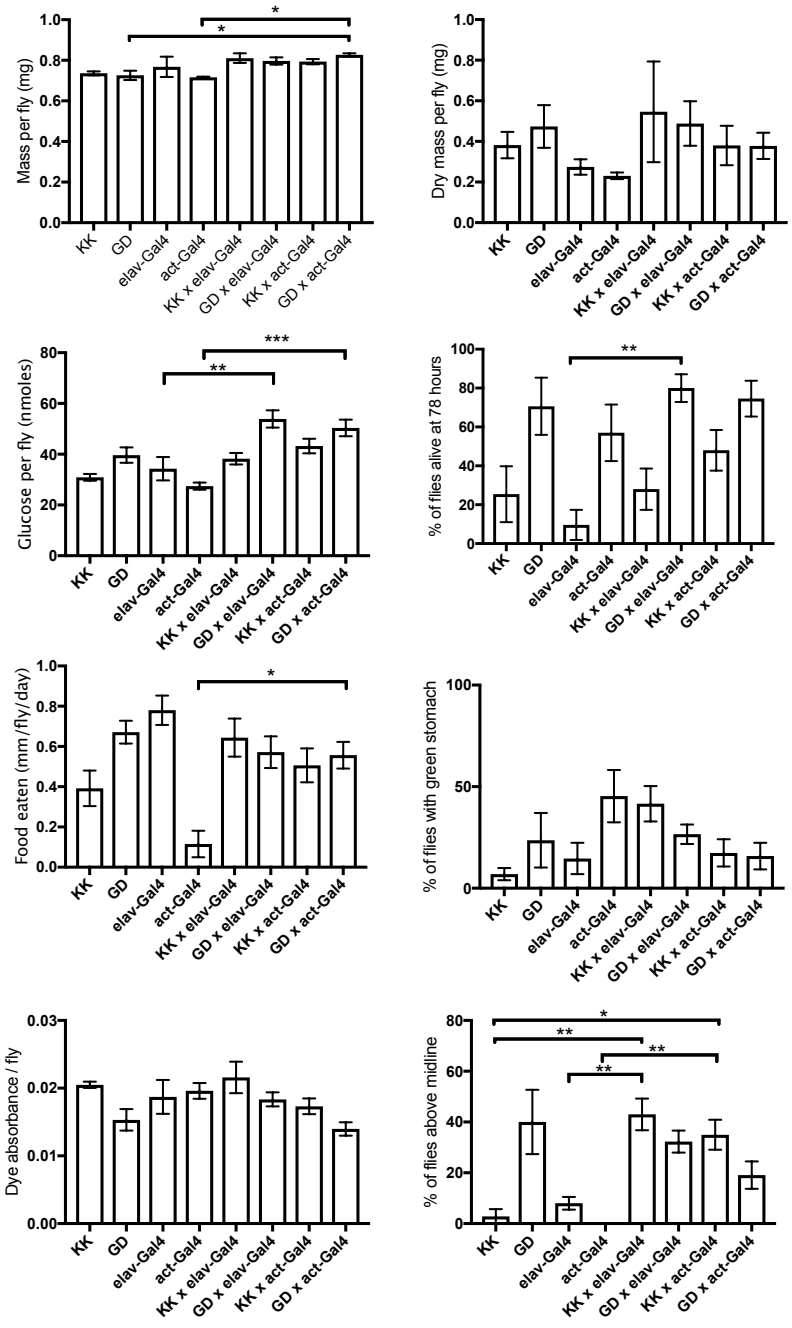


Fig. A.6 The KK, GD, act-Gal4 and elav-Gal4 lines have different phenotypes, so were bred together to create lines for use as experimental controls.
Mean ± SEM is plotted. * p<0.05, ** p<0.01, *** p<0.001, **** p<0.0001.